

Who participates . . . and how?

SYDNEY BRENNER: The discussion goes that the public should know and the public should decide and the public should participate in all the decisions on what research should or should not be pursued. And we are told that these are not decisions for scientists to take alone in laboratories. What I should like to hear spelt out is: how does the public actually come to participate in these decisions? I mean, we're talking about experimental politics rather than experimental biology here.

ANTHONY QUINTON: There's a saying here, isn't there, *sagesse oblige*. The trouble is that the general public is not in a position to make any sort of decision at all until it is told a great deal by scientists. I think one might reasonably say the existing institutional machinery for the public dissemination of socially interesting or socially relevant scientific activity is rather rudimentary and haphazard and improvisatory.

On the other hand, I don't think I'd like

it to be the responsibility of a particular ministry, with civil servants putting in papers. It's just the sort of thing that a Quango might be rather good at doing, where you have a very substantial scientific representation, where scientists would come and talk about their work. A sort of conceptual, theoretical ombudsman in a way.

BBC Radio 3, reprinted in *The Listener*, 26 July 1979.

NIGEL FORMAN, one of the few Members of Parliament with a lively and continuing interest in science and technology, notes the demise of the Commons Select Committee on Science and Technology in this issue (page 443). With the disappearance of one valuable meeting point for scientists, administrators and 'the public', to be replaced by new department-orientated committees in which science is most unlikely to receive as consistent a hearing, it is reasonable to ask whether public understanding of science — or more particularly public understanding of the social issues raised by science — is unacceptably low, as might be inferred from Anthony Quinton's remarks.

What channels are there in Britain for keeping up the links? Obviously there are the printed and broadcast media, with quality ranging from very high to abysmal — but with, of course, relatively few opportunities for two-way communication. There is, sadly, no real British equivalent of the excellent serious American magazines able to devote substantial space to matters of public policy. There is the Royal Society, which has been trying, through its discussion meetings to introduce the occasional large scientific issue to a wide public — but which, too often, flounders on the uncompromisingly technical nature of its lecturers' contributions. There are the British Society for Social Responsibility in Science, the Royal Institution, the newly-formed Science and Technology Association, the Science Policy Foundation and the Council for Science and Society, each trying in distinctive ways to encourage scientists to think beyond the confining walls of their disciplines. There is the Parliamentary and Scientific Committee, an informal meeting point for parliamentarians and top scientists and engineers. There is, of course, a quango already in existence — the Genetic Manipulation Advisory Group, comprised not just of experts but of representatives of the public interest. There is the British Association for the Advancement of Science struggling against formidable odds to inform a wider

public about science — including the social questions raised by science. And there is still some parliamentary select-committee activity, because the House of Lords keeps a rather effective eye on EEC science and environmental policy and may possibly broaden its interests in science and technology in response to the loss of the Commons committee.

But does all this activity add up to a scientific and technological community well aware of social issues? And does it produce a large pool of informed citizens capable of communicating easily with scientists and drawing them into debate on these issues? It does not. Part of the reason is undoubtedly that there are too many organisations covering the same sort of ground, looking to the same sources of financial support, calling on the same pool of people for their core support. But this is not the whole story. When scientists have gone out of their way in the past to discuss their work and its implications with intelligent non-scientists, they have too frequently found the response disappointing, both in terms of numbers who attend in the first place and those who are prepared to make the intellectual effort to stick with a subject after a first encounter.

So, with gloomy experiences in the recent past, there is a temptation to keep further efforts in the field to a low level and restricted to a limited number of people close to the interface. This would be a mistake. There is always scope for new initiatives, but they surely have to be done on a substantial scale and have enough long-term commitment that they do not capsize if one venture turns out to be a flop. The most obvious place in which to experiment is the Royal Society. It has adequate meeting rooms, a sizeable secretariat, connections in high places, easy access to a very wide range of scientists, prestige ... Regular, open meetings, not just in London, designed to bring scientists and non-scientists together are surely not beyond its capabilities and could play a major role in raising Britain's painfully low awareness of science. □



US Congress gives go-ahead to Third World institute

LEGISLATION establishing a new Institute for Scientific and Technological co-operation (ISTC) was approved in both houses of the US Congress last week, following the Senate's agreement to drop its previous opposition to the scheme. Although the budget of the ISTC still has to be agreed, the decision means that the US will be able to present the institute as a definite initiative to the United Nations Conference on Science and Technology for Development in Vienna.

Plans for the institute, first announced by President Carter last year and included in his budget request for 1980 presented to Congress in January, had been "derailed" by the Senate in June. The action took the form of an amendment to the Development Aid Authorisation Bill originally offered by Senator Deconcini, who objected to what he claimed was a "new federal bureaucracy" which would duplicate work already being done elsewhere. Meanwhile, the House of Representatives had approved the ISTC proposal, but recommended against the Administration's wishes, that the Peace Corps be made an autonomous agency within the newly-established international development cooperation agency.

The Administration, with a big political stake in the ISTC proposal, was able to watch the two decisions played off against each other during a series of conference meetings between delegates representing the two houses — the formal procedure by

which differences are ironed out before legislation can be agreed. After three separate meetings, it was agreed to accept the House position approving the ISTC and the Senate position on the peace corps, thus satisfying the Administration on both counts.

The conference report, which was subsequently accepted by the full House and Senate, also stated that the amount of new money it was prepared to authorise for the ISTC should be cut from the \$25 million in the President's request to \$23.75 million. Initially the bulk of the institute's research funds will come from the transfer of research projects already underway within the Agency for International Development (AID). The precise budget will not be settled until both houses agree to foreign aid appropriations next month. The conference also agreed to accept the modifications which the House had previously made to the Administration's proposal for the ISTC. In particular, it has strengthened the role of the governing council so that it will have the right to review (although not, as some would have liked, to veto) all projects costing more than \$200,000 or lasting more than two years. The new institute will have two main functions: to support research into topics of direct relevance to the developing countries and to help stimulate the development of research facilities within the developing countries.

David Dickson

HSE to review safety at Windscale

FOLLOWING the latest incident of a fire at Windscale where five workers suffered low level Cs¹³⁷ contamination, the UK Health and Safety Executive has announced that it will conduct a review of the plant's safety arrangements. A review team consisting of Mr Fen Charlesworth, Senior Deputy Chief Inspector of the Nuclear Installations Inspectorate, Mr Bill Kenney of the Department of the Environment, Mr Stan Gronow from the HSE and a fourth independent specialist to be announced, will investigate safety management, safety surveillance and safety assessment procedures at the reprocessing plant.

The committee has taken the unusual step of announcing ahead of time that it will publish its findings. An HSE official told *Nature* that the large number of incidents at the plant and public concern about the plant's safety have made the HSE want to "let the public see what we are doing about Windscale". According to HSE records the Windscale plant has had 75 incidents since December 1976. Of these, three incidents in addition to the fire are considered major: a still unlocated leak in the B38 silo, a potentially explosive hydrogen build up in the same storage building, and a sump leaking 20,000 curies into the soil earlier this year.

The team will call independent consultants, chosen for specific expertise. They are not expected to include well-known critics of Windscale safety procedures such as Dr Michael Flood and Walt Patterson of Friends of the Earth.

The general investigation has pre-empted a review of the cause of the latest fire leaving numerous questions unanswered. The fire occurred when a fuel disassembly apparatus jammed in a decanning cave. The fire had to be put out manually and radiation escaped contaminating the five workers and causing fallout outside the building. The exact sequence of events is unknown and the questions raised by the incident are central to Windscale safety precautions. These include; how the fire began, why the installed sprinkler system failed to put it out, why the fire burned for as long as 45 minutes, why the fire had to be put out with a fire hose by opening access holes in the roof of the cave, how the radiation escaped, whether the workers were wearing protective clothing and how they got exposed. British Nuclear Fuels has refused to comment on the questions. "We are not prepared to answer questions which will cut right across the internal inquiry presently underway," an official told *Nature*. When complete the report will be released to the local liaison committee.

Joe Schwartz

Pakistan considers uranium enrichment

Pakistan has an ambitious nuclear power programme and, consequently, it is concerned that unless it acquires the capacity to enrich uranium locally, its energy supply will be highly vulnerable.

Two years ago, Canada backed out of the contract it had with Pakistan to provide fuel for KANUPP (the Karachi Nuclear Power Plant which it had already installed) over the issue of whether or not Pakistan was to acquire a reprocessing plant from France. "What do you expect us to do after the experience at KANUPP", said one Pakistani scientist. "If we had not developed the capability of fuelling KANUPP on our own, the 137 MW power plant would have been written-off".

Similarly, a 600 MW reactor being built at Chashma in north Pakistan, received a setback when France backed out of a deal to deliver a reprocessing plant to Pakistan after pressure from the US. Nevertheless, Pakistan's nuclear think-tank foresees many nuclear reactors of either the US light water type or the advanced gas-cooled type being built in the 1980s. Such reactors run on enriched uranium (3% or a little less).

KANUPP, the only nuclear power reactor in the country at present, runs on natural uranium.

With the development of the centrifuge technique, uranium enrichment is no longer beyond the reach of the developing countries. Previously uranium enrichment had relied solely on gaseous diffusion. The situation, however, is now changing. In 1977, even the President of the US approved the gas-centrifuge technique and opposed the construction of a large and costly diffusion enrichment plant.

The centrifuge technique appears to have advantages over diffusion methods where the output of enriched uranium needed is small. A centrifuge complex can be built in stages and a start can be made with only a few hundred centrifuges.

Apart from the advantage that the centrifuge technique presents, Pakistan also possesses two other assets over other developing nations. It has about 500 nuclear scientists/engineers, and sizeable uranium deposits in the Dera Ghazi Khan district.

Azim Kidwai

Comecon tackles the energy crisis

Vera Rich explains how the Eastern bloc is coping with the rising price of oil

Recent oil price rises throughout Eastern Europe, coupled with Romania's unprecedented demand that tourists from other Comecon countries must purchase their petrol for hard currency, at first glance accords ill with Mr Kosygin's remarks at the Comecon summit at the end of June. In fact there is no contradiction. What Mr Kosygin promised for the next Five Year Plan (1981-86) was a 20% increase in energy supplied by the Soviet Union to the bloc, not specifically oil.

Commenting on Kosygin's speech, a TASS reporter quoted Dmitrii Mendeleev, discoverer of the periodic table, that "to use oil or banknotes for heating comes to about the same thing". While the West has based its economy on oil, he continued, the Soviet Union has "harmoniously" developed all fuel and energy branches of the economy: coal, shale, hydroelectricity and nuclear power, as well as the oil and gas fields.

Kosygin's statement of existing Soviet energy supplies to the Comecon allies is impressive: during this Five Year Plan, some 370 million tonnes of oil, 46 million tonnes of oil products, 88 million cubic metres of gas and 64 million kW-h of electricity. But the official communiques of the meeting stressed that from now on the extraction of oil and gas would continue to increase, but would be devoted primarily to the needs of the chemical industry. It seems likely that a major proportion of any increase in energy supplies from the Soviet Union will take the form of electricity. A massive plan to step up generating capacity in the USSR is already under way. From 1980, all increases in the European part of the USSR will be nuclear; east of the Urals, a network of thermal power stations will be brought into the grid, fired by lignite, of which there are abundant local supplies.

All Comecon countries, with the exception of the Soviet Union, are now net importers of petroleum. Even Romania, whose role as a former major oil producer is commemorated by an oil rig in the state coat of arms, has to import a certain proportion from the hard currency market. Although energy is one of the specific sectors selected for the "long-term target-orientated" Comecon integration programmes approved in 1976 (for the period 1975-1990), Soviet oil prices to the other Comecon countries are calculated from the average world price over the last five years and reassessed each year. Moreover, Soviet-exported oil accounts for only 80% of the other members' demands: the rest they must purchase on the world market.

Furthermore, the long-term availability of Soviet oil appears to be in some doubt. All official sources unequivocally



Russian pipeline: the oil will keep flowing, say Soviet officials

repudiate the CIA predictions of an impending Soviet oil crisis. Shortfalls are attributed to the technical difficulties of bringing the East Siberian fields on-stream, and the consequent depletion of the West Siberian fields more rapidly than planned. Occasionally, too, "anti-social manifestations" among the oil workers may be blamed for a temporary deficit (*Bakiskii Rabochii* 1 June, 1979).

Nevertheless, nuclear power is seen as the mainstay of the future energy needs of the bloc, and the construction of nuclear power stations is a top priority in all member countries except Poland. Here the still abundant coal reserves make it possible to defer the introduction of nuclear power on a commercial scale, and the one nuclear power station so far scheduled (a joint venture with the Soviet Union) is seen, in the words of one energy expert, as "a pilot project for the Polish situation".

Socialist planning, it appears, leaves no room for an environmental lobby against nuclear power. A recent Hungarian publication stresses that "in the philosophy of socialism, environmental detriments may present no limitations to the welfare of society. The growing demands of society must be satisfied, since failing to do so would contravene the basic principles of socialism".

So far, only the dissident "Document 22" of the Charter-77 movement has dared to question the validity of the nuclear solution. By 1990, according to the latest agreement, the nuclear generating capacity of Comecon will have been raised from the present 15-18 million kW to 150 million kW, 37 million kW of which will be sited outside the Soviet Union.

Meanwhile, in spite of the agreed integration programme, the search for home-based fuel sources continues throughout the non-Soviet members of the bloc. Poland has a long-standing programme of coal mine modernisation, including a sophisticated research programme into the causes of mine-gas explosions. A country-wide geological

survey for oil and gas resources has produced an impressive set of maps of the various strata, and is now proceeding to the field-work stage. No detailed interim results are available, although one geologist connected with the survey told *Nature*, last autumn that there were "positive indications". The survey is being continued offshore by an international consortium of Poland, East Germany and the Soviet Union.

Romania, too, is looking offshore. The new draft programme for research and development to be submitted to the forthcoming congress of the Romanian Communist Party stresses the need for offshore exploration and exploitation. Not only will this include mineral wealth, but the possibility of building a pilot wave-powered generating station is also under consideration. Full use of the hydroelectric potential of the Danube is a target for 1990, lignite and shale mining will be intensified, and a number of more exotic schemes — "biogas", the extraction of hydrogen from water, and the use of geothermal energy — are to be given serious consideration. By 1990 it is intended that the country should be self-supporting in energy, while the average energy consumed per 1000 lei of industrial output is intended to fall by at least 40%.

Czechoslovakia is looking closely at the home: modern flats, said the Director of the State Power Directorate recently, consume about 50% more heat than the norm, due to poor construction and installation. Yugoslavia (not a member of Comecon but closely linked with the bloc) has already decided not to build any further oil-burning power stations and to upgrade its neglected coal and lignite resources. Finland, likewise partially dependent on Soviet oil, is seriously considering the energy potentialities of scrub timber grown on marginal land. Hungary is showing a renewed interest in coal, and the latest updating of the UK-Hungarian Joint Commission on Technological Co-operation agreed last

month includes a new cooperation project on coal technology. Bulgaria, while so far urging no major changes, is cautiously contributing to the energy debate by publishing statistics of kW-h of lighting saved this year by "summer time".

Meanwhile, East Germany, although claiming that, in spite of the harsh winter conditions, planned production of electric power was "achieved as a whole after the first six months", did not allow the temporary setback to the economy to be

entirely attributed to meteorological conditions. In June, a result of "shortcomings in management, especially in the ministerial sector for coal and power", Klaus Seibold, the Minister of Energy, was dismissed from office. □

Spanish scientists have to fight for survival

Research in Spain is suffering from a lack of direction and funds. **M Vicente**, a Spanish biologist, reports on a recent meeting of scientists to discuss their plight and **Emilio Munoz** (below) reports on one ray of hope — the setting up of a new research ministry.

Uncertainty of direction, lack of promotion opportunities and no guarantee of long-term finance: these are the crucial problems that bedevil scientific research in Spain. They were given a public airing last month when biology and biomedical researchers met under the auspices of the CSIC (Consejo Superior de Investigaciones Científicas). The scientists' intention was to test the reaction of the newly-created Ministry of Universities and Research (see below) to their complaints, and what emerged from the researchers amounted to a cry for survival.

The meeting had no difficulty agreeing that research policies were needed related to Spain's particular requirements. But they could not decide whether this was best approached through basic or applied research. One speaker said biology in Spain had concentrated on basic research for the past 15 years and all applied approaches had been treated as secondary. It was not surprising, he added, that there were no links between society and research and that, as a result, society was indifferent to

the present scientific crisis. Other speakers defended basic research and argued that they were doing enough for society by teaching young scientists in the universities.

The second major theme of the meeting was that the country's scientific researchers were not only too few in number but also too old. The average age of Spain's scientific community was said to be around 48 and this restricted the job opportunities for young scientists. The number of young PhDs nearly doubled every four years but only one in five had a job related to his specialisation. Several causes were given for this situation, all with the same roots: the absence of long-term planning and the lack of interest in research.

Spain's scientific development during the 1960s happened at the wrong time and went in the wrong direction. The economy caught up with other developed countries, but largely because of imported foreign technology. A 10%-a-year increase in public sector research staff was intended to accompany economic growth but never

came into effect. Researchers in the private sector are even fewer and their work is limited to quality control. All these problems have now been aggravated by economic difficulties, so that Spanish research is coming of age at the worst possible time.

For scientists, the finance situation could be the most worrying of all. Because of its lack of social impact, research is not in the official list of priorities. Half way through the present year, research groups are spending the money assigned to them last year. No official programme of grants has been issued yet for next year. Unless urgent measures are taken, and experience of Spanish bureaucracy makes it unlikely that they will be, most scientists will have to stop their work early in 1980. Even strong groups, such as those in the Molecular Biology Centre in Madrid.

A secret three-year government plan to finance research was drafted a year ago. Its text is not known, raising suspicions among scientists that it will channel money to the increasingly overcrowded universities in which research has been so difficult that it tends to be the activity of a few dilettantes. They are afraid that a decision to transfer them to academic jobs will irreversibly damage their research activities. □

Will the new research ministry fill the vacuum?

SPAIN has established a Ministry of Universities and Research — a move which has been welcomed by the scientific community. Scientists have felt that the lack of an effective government science policy has been paralysing efforts to build a new structure for research in Spain around the re-constituted Consejo Superior de Investigaciones Científicas (CSIC), and they hope the new ministry will fill the vacuum.

The Ministry of Universities and Research has been created by splitting the functions of the former Ministry of Education and Science, in which science had become submerged in the complicated problems of education. However, there are still fears that hopes for specific research proposals may again be disappointed. The government has not yet presented its research programme, so the reasons for the creation of the new ministry are unclear. And there is still a lack of communication between the scientific community and society. No national debate has taken place in parliament on this subject, although various sectors in the country have been

demanding such a debate since 1977.

The need for a new ministry was established more clearly when the new statute of the CSIC was first put into practice. It was published in January 1978 and came into force the following June. It included a new system of government for the CSIC with the participation of collegiate bodies as advisory organs, and responsibility for the economic and scientific fields through the Scientific Council and Economic Council. This made real change in the CSIC a possibility. As with all the Spanish organisations during the time of the autocratic regime, CSIC had been conducted without any clear intent. However, the councils suffered many difficulties during their first months, due to a lack of scientific policy. The absence of any sign of a specific scientific policy has prevented a rational reorganisation of CSIC.

The reorganisation attempted by the CSIC so far has been mainly bureaucratic and administrative and is thus running the risk of becoming neither lively, dynamic nor relevant. However the Scientific

Council of CSIC has made imaginative efforts to amend its organisation. A document on the general trends of scientific policy in CSIC has been elaborated, which considers as fundamental that the activity of CSIC be adjusted to the needs of Spanish society and to the manpower and material resources of CSIC itself. The budget for 1979 allows 100 pta. (less than US \$2) per day per research worker for expenses of nonrecoverable material — and this represents a significant improvement over 1978!

Within the limited possibilities offered, CSIC has strived to overcome one of its main criticisms — its isolation from other research institutions. It is stressing agreements with universities and research institutions with the ultimate goal of resolving the existing academic power struggles. The Scientific Council of CSIC is trying to cooperate with the new Ministry of Universities and Research and recently asked for an interview with the minister with the aim of establishing contacts. □

news in brief

Five French unions oppose CNRS reorganisation: Plans by the Giscard d'Estaing government to reorganise the Centre National de la Recherche Scientifique, France's national research council, have been met by opposition from five of the trade unions involved, according to recent reports in *Le Monde*. Mr Guy Hermier, a deputy from the Bouches-du-Rhône constituency, vice president of the national assembly and a member of the political bureau of the Communist Party, has called for a national debate on the future of the CNRS to be preceded by expensive consultations with the scientific community. In a letter to the Prime Minister, Hermier accused the government of giving up control of research to industry.

The reorganisation plan proposes a reduction in numbers of representatives on national CNRS committees, the addition of representatives from industry and the exclusion of technicians and administrators who are currently represented. The streamlining of the council and the involvement of representatives from outside the inner circles of the CNRS is seen by the government as an important move to decentralise research and to make results more accessible to the needs of French industry.

BSSRS criticises HSE on MMMFs: The Work Hazards Group of the British Society for Social Responsibility in Science has questioned a recent UK report on the hazards of mineral fibres used for insulation (19 July page 183). Although welcoming the tripartite (industry-trade union-government) report from the UK Health and Safety Executive which warns that "it would be prudent to regard these fibres with suspicion", the Work Hazards Group raises "considerable" objections to some of its conclusions. In particular the report is criticised for underplaying the risks to the domestic householder, "minimising the risk of bronchitis due to high dust levels and inadequate dust-exposure and emission controls. Citing the known dangers of respirable dusts and fibres below 3.5 microns in diameter to the air sacs of the lungs, the Hazards Group criticises the classification of mineral dusts as nuisance dusts. "We know from asbestos that the small fibres cause pneumoconiosis and mesotheliomas" says Dave Hayes of the Hazards Group. "And all mineral insulation materials contain these respirable fibres." More generally, the group is critical of the Department of Energy for promoting mineral fibres for the do-it-yourself home insulation market without "due regard for the hazards involved".

The Work Hazards Group recommends a sharp cutback in the production of small fibres (which it says are unnecessary for insulation), strict controls on home and industrial use of dust-causing insulation materials, notification to home owners that a "good" respirator should be used and that trade unions should keep their own medical records for workers at risk. Recommended substitute materials are insulation board, vermiculite and cellulose newsprint used with a good respirator. *Mineral Fibres. British Society for Social Responsibility in Science, 9 Poland St, London W1, 13p.*

Three Mile Island citizens panel dissolved: The shortest governmental advisory committee in many years was dissolved on 25 July without holding a single meeting. Four weeks ago, the Kemeny Commission which is investigating the accident at the Three Mile Island nuclear plant in Pennsylvania, appointed a nine member citizens advisory board to give it some "outside input". The Commission, apparently, got more than it bargained for when the panel demanded copies of all staff memos, draft reports and work plans, access to all federal and industry documents subpoenaed by the Commission and depositions taken by the legal staff. In addition the panel demanded the opportunity to review lists of witnesses, to demand additional witnesses and to criticise and review the final report before it went to the White House. The Commission had envisaged the citizens committee taking a more passive role by answering nine questions put to it. "They wanted to be inside and we felt that was

unacceptable" said Barbara Jorgenson, press officer of the Commission. At a meeting on 5 August the Commission voted to circulate the questions to the groups represented by the original nine panel members and to solicit suggestions and answers from an additional 200 public interest groups.

Life or death of an observatory: The Centre for Geodynamics and Astronomy, near Grasse in Southern France, was planned to co-ordinate all observational activities in the fields of positional astronomy and geophysics. The first interferometric star observations were obtained last year using a 30m baseline Michelson interferometer. But only 6.5 km away, a quarry and stone-crushing operation has created dust that will diffuse the night light from the coastal towns which will make very faint objects completely invisible. The large Schmidt optical telescope, installed last year, now might not be able to observe more than just a few percent of the interesting stars. Another problem is vibration from quarrying activities which disturbs high precision optical and long-baseline interferometric measurements.

Appeals to the local authorities have failed and the CERGA has petitioned the appropriate ministry. But will the case be taken up in time, before the site is permanently ruined, by government bureaucracy?

From the staff of la Recherche



USSR higher education shows sharp regional discrepancies: A working note issued by the German Rectors Conference shows that higher education in the Western regions of the USSR continues to outstrip that in the other republics. The western republics lead in higher education with 200 students per 10,000 inhabitants. This is followed by the Baltic republics with 185, the Caucasian republics with 175 and the central Asian republics with 159.

The rate of increase of higher education also favours the western region with the number of students increasing relative to the population by a factor of 2.58 since 1966 while the Caucasian republics brought up the rear with an increase of only 1.06. The Rectors conclude that the discrepancies are small "considering that the Central Asian Republics can be regarded as developing areas." *CRE-INFORMATION No. 46. Standing Conference of Rectors and Vice-Chancellors of the European Universities.*

Solar activist to head research institute: The US Department of Energy has appointed Denis Hayes, a prominent leader of the environmentalist lobby and currently a senior research fellow at the Worldwatch Institute in Washington, as new director of the Solar Energy Research Institute (SERI) in Golden, Colorado. Hayes, one of the main organisers of last year's international Sunday, will at the age of 34, become the youngest director of a national research laboratory. He succeeds the current director, Dr Paul Rappaport, who has headed the institute since it opened, and who will remain at the institute as distinguished scientist.

"My goal is to make SERI the most exciting and effective energy research organisation in the world," Hayes said last week. "The institute has a staff of first-rate analysts already thinking through the strategies and research requirements. Now its time to get those programmes moving".

Yoshimura has to leave Britain: Dr Tetz Yoshimura, the Japanese mathematician, has lost his fight to stay in Britain. He left for the continent earlier this week, a few hours ahead of a deportation order. Dr Yoshimura told *Nature* that he hopes to wait in "a country with less stringent immigration laws" for an invitation to a new academic appointment "somewhere in the third world".



'Always on tap, never on top'

Government scientists this week called off their industrial action after settlement of their prolonged pay dispute. **Joe Schwartz** reports on the scientists' attitudes towards the strike whose success took both union and management by surprise

"WE'VE ALWAYS been known as a non-militant union. I guess the Civil Service Department thought we would stay that way", declared a local committee member during the strike by the Institution of Professional Civil Servants. The 104,000 members of the IPCS amazed the public and fellow trade unionist with five weeks of action to back up their pay claim against the government.

To outsiders, the scientists and technologists in the UK's Civil Service are unlikely candidates for militant industrial action. Their union, one of seven in the Civil Service, only affiliated to the Trades Union Congress in 1976 by a vote of 57,000 to 38,000; the minority argued that trade union affiliation was "unprofessional". But strength of feeling built up through a series of skirmishes with their Civil Service bosses dating back to 1974. One member voiced the general feeling: "People think that civil servants are advisers to ministers. We're not. We're ordinary scientists and technologists who are sick and tired of being taken for a ride. Whenever there is money to be saved, the government tries to take it from the scientists".

With this growth of feeling, the only question earlier this year was how the IPCS's 25-member elected national executive and its full-time office staff of 30 would organise themselves for the annual pay talks with the Civil Service Department in April. "In fact, we organised very poorly," says Bill Brett, assistant general secretary of IPCS. Brett, a trade union organiser for 15 years in four different unions (including six years with Clive Jenkins of the Association of Scientific Technical and Managerial Staffs), expected that it

would take a dramatic move on the part of the CSD to galvanise the membership. The CSD obliged. They refused to recognise a promise given the previous year that scientists' pay rises would be linked to civil servants' in the administrative grades. After granting 25%-45% to the administrators, the CSD offered only 18% to the scientists.

Scientists and technologists saw it as the culmination of what some alleged was a campaign in the CSD to downgrade the scientists. Tired of being "always on tap and never on top", they met in conference and instructed their leadership to organise industrial action.

The union drew up a battle plan of half-day strikes, working to rule, one-day demonstrations and the critical selective strikes that were to bring work in many government laboratories and defence establishments to a halt. The plan was put to the members at a series of meetings throughout the country. "We spoke to

50,000 members at about 25 evening meetings of from 500-2,000 people over a period of two weeks", says Brett. "We received an overwhelming mandate to go ahead."

For scientists at the Central Veterinary Laboratory in Weybridge, the strike call gave them the chance to use the experience gained two years earlier when an "informational picket" turned into a fully-fledged strike in support of a claim for an outer London pay allowance that had been granted to neighbouring laboratories.

The Weybridge laboratory does research on animal diseases. It is responsible for reviewing the quality and effectiveness of commercially produced drugs and it operates a P4 containment facility for diagnostic work on rabies. Work at the laboratory was brought to a halt when the local IPCS committee pulled out two of its supervisors from the laboratory's steam room. Members of the Transport and General Workers Union refused to work without supervision, and the laboratory's five high pressure boilers were shut down, cutting off heating, autoclaves and distilled water facilities. Eight more members were called out and a picket line set up which lorry drivers and post office workers refused to cross, thus halting the delivery of supplies and samples to the



Digging in: Striking scientists at the Laboratory of the Government Chemist in London combined three weeks of picketing with some voluntary community work, clearing up allotments for old people. Leaving one person on the picket line, the strikers kept themselves busy on a daily work brigade.

laboratory.

The picket was not total, however. IPCS members decided they would give exemptions in cases of safety and public health. Thus, at Weybridge, operation of the P4 containment facility was guaranteed, as were deliveries of liquid nitrogen to maintain the laboratory's reference cultures of viruses, parasites and bacteria. Applications for exemptions went through the elected local committees; doubtful cases were passed on to the union executive committee.

At the Laboratory of the Government Chemist in London, a management request to allow gas deliveries through was refused. "All they wanted to do was to continue as usual. There was no question of safety", explained Arnold Pinder, secretary of the LGC section. Chemists at the LGC found that they too tried at first to carry on as usual. "Your first inclination is to keep working. You don't really want to strike and you look for ways to get around it. We had no chemicals, no water and no service. So we looked for other things to do", says Pinder.

A significant feature of the IPCS strike was the solidarity shown within the heterogeneous membership. Scientists stayed out on strike for three further weeks in support of the technologists after their own pay claim was settled. Again the Civil Service Department appears to have underestimated its adversary. "It was so obvious", says an LGC member. "First they split off the union with the least militant reputation and the most split membership. Then they tried to divide the scientists from the technologists, and then they tried to settle with the younger scientists at the expense of senior staff by offering the first two steps of the rise only." The tactic unified rather than divided the membership. "The CSD did at a stroke what the IPCS has been trying to do for years."

The IPCS action achieved a number of specific results and one important general result. The scientists are to set the same pay rise as their administrative colleagues this year. The technologists, however, feel they have a basic principal to settle about their status and comparability. This is now to go to a conciliation procedure.

But most importantly, IPCS members have realised that they can exercise effective industrial muscle. It has surprised both union and management. Researchers are reluctant to strike and have difficulty in seeing how withdrawal of their labour can be effective. Scientists in the IPCS have learned how to do it. This will change the complexion of all future pay negotiations. As Arnold Pinder put it: "We don't like to strike. But no one will listen to you unless you strike. You have to show your power before they'll listen." □

Ireland, land of priests . . . and engineers

Robin McKie compares the high status of the Irish engineer with the low status of engineers in Britain

THE realisation that Britain's manufacturing industry is far behind our competitors in terms of flexibility, sophistication and output has been forced upon the nation in the wake of its decline as a major commercial and financial force. And certainly any understanding of the problems which afflict the country's technological capabilities requires some form of detailed comparison with the performances of our closest competitors.

Indeed, in the case of our nearest English-speaking neighbours, the Irish, the contrast is striking — and decidedly unflattering. Last year the gross national product of Ireland expanded by 5.5%, almost twice the European average; manufacturing investment increased by 20% and manufactured exports rose by more than 20%. This year, industrial output is expected to rise by 11% — compared with 2.25% in the UK! This has been achieved by attracting a vast amount of foreign investment, mainly by permitting companies to make tax-free export profits until 1990.

It may allow large amounts of money to slip out of Ireland, but it has also succeeded in attracting more than 750 manufacturers — including major microelectronics, drug and oil companies — in the past 20 years, produced fixed assets worth £1500 million and in 1977 alone saw the creation of 24,000 new jobs.

In a sense, these measures were forced upon Ireland — a nation which has never been rich or prosperous. However, it is hoped that eventually the resulting financial regeneration and accompanying technological infrastructures will help transform the country into a major independent industrial force. And certainly "the Irish miracle" represents a remarkable achievement and indicates the country's critical recognition of the importance of high technology.

This last factor reflects the influence of scientists and engineers in Ireland and their importance within industry and government. In Britain, the engineer and technologist is all too often relegated to a minor, supervisory role inside a company, while his Irish counterpart wallows in high status and recognition.

As Professor P. Leahy, dean of engineering at University College Dublin, put it: "Over here, an engineer is looked upon as someone like a doctor or a priest".



Religious shrine in Dublin — also a city of developing industry

And a more striking illustration is provided by the fact that the chief executives of Ireland's broadcasting corporation, transport system, electricity board and many of its investment banks were all originally engineers.

This obvious high achievement rate has had the effect of encouraging the very best students to take up engineering. For instance, in 1977, 15% of entrants to University College Dublin's engineering school had scored 30 or more points in the national school leaving examinations of Ireland. This compared with 7% of those going into medicine, 0.72% into law, 0.41% into science and 0.19% into arts.

This is a remarkable contrast to Britain where often the lowest qualified students form the largest proportion of engineering students. For instance in the period 1973-75, a total of 42% of students taking engineering at universities had worse than the equivalent of three Cs at A level. In science the proportion was 34%, in medicine 26%, and in social, administration and business studies, only 24%.

For the Irish, this concentration of talent has resulted in engineering schools becoming the prime recruiting grounds for chartered accountancy firms and investment banks who sell employment as a major route into management. (Indeed, many of the academic staff at UCD's faculty of commerce also have engineering first degrees.) And in this way an appreciation of technological matters becomes embedded in the fabric of Irish management and industry.

However, the reason for the original high status of engineers in Ireland is less easily defined, although it is closely connected with the industrial growth of the country and related educational routes.

In Britain, a rapid industrial revolution

led to a quick demand for engineers who could not be supplied in sufficient numbers by the few existing university departments. So most rose from shop floor workers and craftsmen to become engineers — a process that carried little prestige. But in Ireland — which generally escaped the rigours of industrial revolution — most engineers were civil engineers, with university degrees, who provided roads, bridges and other services for the community and carried a corresponding high status.

Indeed the pervading nature of the differing prestiges of Irish and British engineers is explained by the fact that in 1958, 95% of those achieving engineering status in Ireland did so through university

degrees. In Britain, the proportion was still only 39%, the rest achieving chartered status via part-time exams set by the professional engineering institutions. Only in recent years has Britain approached Irish levels for university educated engineers. For instance, in 1974, the figure had reached 90% in the UK, and 93% in Ireland.

Thus a picture can be seen of an original, well-established prestige for engineers in Ireland, leading constantly to better students and eventually to an even higher status. And certainly a thorough confidence in its engineers and technologists, combined with its recent industrial re-birth, indicates a strong financial future for Ireland.

It is then all the more vital to have the right quality of technical manpower to sustain this industrial growth. As Matt O'Donovan, of the Institute of Engineers of Ireland, stated: "It is important that we are training not just for existing industry but to initiate new industry".

It is this last capability that is particularly absent in Britain at present and which must be restored through the proposals of the Finniston committee of inquiry into the UK's manufacturing industry. However it is hard to see how it can find a way to cut through effectively the damaging spiral of engineers' low status, producing poorer levels of student intakes and eventual graduates, and resulting in even lower status for engineers. □

How industry lost an enthusiastic engineer

Robin McKie tells how one technologist's enthusiasm and ambition turned to disillusionment

For Barry Francis, engineering was his only choice of career since the age of nine. It seemed an obvious selection for a youth fired by an early enthusiasm for constructing crystal radio sets and who eventually ended up building hi-fi systems as a hobby.

So after he took O and A levels at Goring Hall, Sussex, he entered Middlesex Polytechnic for his main engineering course and later transferred to Manchester University for a postgraduate diploma in semiconductor electronics. Then in 1969, Barry moved to industry where he began work as a research engineer for Hawker Siddeley at Hatfield.

And almost immediately disillusion set in. "I discovered rather quickly that I would not be pushing back the boundaries of discovery every week. Instead I found myself working on projects that would not come to fruition for 10 to 15 years", he said.

Although the work — designing special instrument displays for vertical take-off jets — provided reasonably fulfilling employment in a limited fashion, it was also clear that the aircraft construction industry was far from healthy. So, with no immediate job satisfaction and no obvious career path, Barry left Hawker Siddeley after only 18 months and moved into mainstream engineering.

At Control Systems, Uxbridge, he was appointed a senior research engineer jointly in charge of a project team of 15 technicians, programmers and engineers who were designing a small business computer to control payroll, invoicing and other office chores for companies of up to 100 employees.

Sadly, although the work again was reasonably fulfilling, disillusion quickly set in once more. A particular problem lay with the firm's controlling company which eventually seemed to lose interest in the project and the team found itself

increasingly pushed lower and lower on lists of priorities.

However, more importantly, Barry believes that although the project was a reasonably modest undertaking, it required more in terms of human resources than the company could then provide.

These experiences, plus involvement as a writer for engineering journals, led 34-year-old Barry to the view that a major flaw in UK industry is its amateurishness. This is not the particular fault of its exponents but there can be little doubt, he feels, that British scientists, technologists and engineers have a poor feeling for industry.

And a major portion of the blame must lie with UK universities and polytechnics. "My own education did not in any way prepare me for industry's requirements. There was no design, management or marketing training — and these are crucial for you cannot just build a piece of equipment in isolation without knowing if there is a need or a sales potential for it. Really, at university and college, we were just being prepared as glorified technicians".

Generally Britain can be seen as a nation that cobbles things together to produce a solution under pressure. Often budget constraints can help to find more effective production methods but all too frequently initiatives are lost to other countries who are capable of more adventurous and stylish marketing and design systems.

After his two depressing experiences in industry, Barry left to be a writer and associate editor for various engineering magazines before later moving on to his present occupation as a public relations consultant. And from his story it is hard to avoid the view that British industry, because of its very make-up, lost someone of ambition and drive who is now unlikely to return. "In a sense I cannot move back at the moment without having to make a big financial sacrifice. In fact, if I went into

industry again, I reckon my income would be halved."

This view has often been endorsed by various studies of engineers' conditions in industry. For instance, Peter Lawrence, of Southampton University's engineering department, recently produced a paper which showed that salaries in the UK were greatly below the levels of those in Germany. In 1976, the average salary for an engineer in the UK was £5510 while in Germany it was £12,119 — and at that time the cost of living was only 36.4% higher than in Britain.

Similar discrepancies in awards exist in most other European countries, including Ireland. As Professor Scanlon, head of University College Dublin's electrical engineering department put it: "I can give more to a PhD student in Ireland than industry gives as a salary in the UK".

Of course poor salaries reflect the low status of engineers in Britain — a view backed by Barry Francis. "Not enough bright kids are going into engineering", he said. "They see lawyers and doctors living well and decide to follow them."

"However once you have put this right and have raised the prestige of engineers, the bright children will start coming in, and then we'll eventually get better British products and an improved industrial performance".

But even if conditions were improved, there would still be difficulties in attracting the return of former engineers such as Barry Francis. All too often, these people find their conditions and trainings badly out of date and discover there is no system for updating engineers in recent technological developments. Proposals — such as one-month university courses — are now being considered by the Finniston committee and the Science Research Council is also set to launch a trial project on these lines.

Only when a total package of proposals on these lines is implemented can Britain then hope to attract the talent and ability needed to restore its technological and industrial strength. □

Sadly missed: a key contributor to science policy

IT IS not just those of us that had the good fortune to serve on the Select Committee on Science and Technology who regret its demise. There must be many others outside the House of Commons who recognise the valuable contribution which it made to a wider understanding within Parliament of certain aspects of science and technology and to the quality of some government decision-making.

Since its establishment in January 1967, the Select Committee investigated a wide spectrum of subjects, ranging from nuclear power to coastal pollution and from carbon fibres to computers. It was chaired and guided by two distinguished politicians, Arthur Palmer from 1967-70 and again from 1974-79 and the late Airey Neave from 1970-74. There was also a sensible degree of continuity in its relatively small but enthusiastic membership.

It is hard to say exactly how much influence was wielded by the committee, because government departments are not in the habit of revealing to the public the various stages of their decision-making procedures and in our constitution ministers claim sole responsibility for the decisions which are actually taken. However, the record shows that the committee's 1969 report on the UK nuclear power industry which proposed the establishment of an Atomic Energy Board, was eventually followed by the Atomic Energy Authority Act 1971; that the 1971 report on population policy which called for the establishment in government of a special population office, was followed by the creation of the Population Panel; and that the 1974 report on energy conservation was eventually followed in December 1977 by a substantial, if belated, package of energy conservation measures which included some of the recommendations made by the committee.

In other cases, although the government of the day initially reacted in a defensive, even hostile, way to some of the committee's recommendations, subsequent events proved them right or led



Nigel Forman MP mourns the passing of the House of Commons Select Committee on Science and Technology

government to reconsider its earlier policy response. One example is provided by the way in which the last government's initially negative response to the committee's advocacy of alternative sources of energy in its 1977 report was eventually transformed into a more positive view of the need at least to explore the technical possibilities in this area and to do so with the support of adequate R&D funding. Another example is the way in which the Departments of Environment and Transport eventually came to accept the need for the Planning and Transport Research Advisory Committee to include in its terms of reference and its annual report a specific account of R&D work on advanced ground transport, something which had been rejected by government in its initial response of August 1974.

As for the work in progress at the time of the committee's demise, it is impossible to tell yet whether or not the uncompleted investigations into recombinant DNA research and developments in small passenger car engine technology will be pursued by the relevant "departmental" Select Committees which are to take its place. With any luck this will happen and a great deal of useful evidence will not have been wasted. Both these investigations exemplify the tendency of the Committee latterly to balance its work on short-term subjects — such as filament and discharge lamps or the lessons of the Eleni V disaster — with more important long-term work on emerging technologies for the future which may have significant benefits and costs for our society in the 1980s and beyond.

Now that the committee is no longer in

existence, it is possible to form a clearer assessment of just what will be missed by Parliament, the press and the general public. Members of the House of Commons will miss the opportunities which the committee provided for consideration of important scientific and technological issues which by dint of their scope and complexity, are bound to transcend departmental boundaries. Recombinant DNA research and the wide-ranging implications of innovation, R&D in Japanese science-based industry are two typical and important examples.

Journalists will miss the platform which the committee provided for the public airing of important questions which are not otherwise debated with any frequency or in any depth in the normal proceedings of Parliament, whether in the Chamber or in legislative Standing Committees. The general public will be deprived of a useful Parliamentary mechanism for encouraging at least some of their elected representatives to look beyond the inevitably short-term horizons of contemporary politics. This last consequence is perhaps the most severe blow, since it is now more important than ever to preserve and, if necessary, create institutional mechanisms which allow arguments about social time preference and encourage the broadest possible cost-benefit analysis to be undertaken across the whole spectrum of contemporary science and technology.

Of course, in many ways the influence of the committee will live on through its *alumni*, like Neil Macfarlane, the junior minister with responsibility for science, who now find themselves in influential positions in government, and through its example in seeking to tackle a wide range of scientific and technological issues which will not go away just because it no longer exists. Furthermore, it is not as if there are not other laudable institutions of these issues. There is the well-established Parliamentary and Scientific Committee which has played a valuable role since its inception soon after the Second World War. There is also the Council for Science and Society which deserves to find continued support for the useful work which it does in attempting to bring scientists and laymen into a state of greater mutual understanding.

However, the fact remains that the peculiar dual role of the Select Committee as an educator of Members of Parliament and an investigator of issues which would not otherwise be the subject of Parliamentary scrutiny, will be difficult to ascribe to any one of the "departmental" Select Committees within the new system recently endorsed by the House of Commons. It must, therefore, become an important objective of those concerned with these matters in the new Parliament to meet this requirement by other means. □

Select committee considers GMAG

THE final report of the Select Committee on Science and Technology, released last week, is the result of the committee's recent investigations into the "public safety and public policy" issues of recombinant DNA.

In a brief summing-up of the evidence taken, the report expresses concern at the way in which the Genetic Manipulation Advisory Group handles notifications of industry's intentions to work with recombinant DNA. Under present arrangements, members of GMAG with potential industrial interests and those who have not signed a confidentiality agreement, withdraw from discussion of

notifications by industry. The Select Committee disapproves that an advisory body "set up to safeguard both workers and the general public, should have apparently first and second class members". The report also expresses "surprise" that the Department of Education and Science (DES), not the Department of Health and Social Security (DHSS) is responsible for GMAG. It is concerned that the strong position of UK researchers and companies should not be hampered by unnecessary regulation. International safety standards also "appear to be desirable", it says.

correspondence

Mongolian science

SIR, — In connection with the interview with me, by Vera Rich in *Nature* (28 June, page 754) I would like to ask you to publish the following explanation.

When Vera Rich visited me in Warsaw in October 1978, she was astonished to learn from me that there is a Museum and Palaeontological Laboratory in Ulan Bator, with a group of competent Palaeontologists, some of a very high standard.

I was shocked when I read what she published in *Nature* some 9 months later — an article with ironical remarks about the scientific relationships between the CMEA countries (Comecon in Vera Rich's wording) and about Soviet science; and containing misleading information, for example that R. Barsbold's article (28 June, page 792) is based on materials from the Polish-Mongolian expeditions, while in fact it describes the specimens collected by the Soviet-Mongolian expeditions.

The unacceptable side of her article is the perjorative and condescending tone in which she writes about Mongolian science and scientists. As her article is written in a form of an interview with me, it makes me look disloyal to the people with whom I have been co-operating for years. I have never described my Mongolian colleagues in a derogatory way. What I have done was the opposite, but this apparently did not fit Vera Rich's ideas about Mongolia.

As a result of this interview co-operation between Polish and Mongolian palaeontology may be impaired. Is this really the aim of *Nature's* policy?

Yours faithfully,

ZOFIA KIELAN-JAWOROWSKA

Academy of Sciences Institute of Palaeobiology, Warsaw, Poland.

Readers concerned that we may have seriously slighted Mongolian scientists are urged to read Vera Rich's original article. Ed.

Alpha-fetoprotein screening on a US regional basis

SIR — The future of pregnancy alpha-fetoprotein (AFP) serum screening in the United States (5 July, page 6) is dependent upon decisions from the political as well as the scientific communities. Intense pressure has been applied to the Food and Drug Administration (FDA) to restrict test reagent kit licensure to carefully monitored regional programme. Few such programmes currently exist, due to lack of financial support from major funding sources such as the Department of Health, Education and Welfare, thus placing the FDA in a difficult position.

A number of private and commercial laboratories in this country have produced their own reagents and now offer the AFP test for pregnancy. Such testing is legal and is not covered by FDA restrictions. In the absence of coordinated action by federal funding and licensing agents such testing will proliferate and its quality will be unpredictable. Equally

important, the ability to interpret and follow up properly on test results will be generally poor, especially in the absence of an integrated patient service programme. These issues have been extensively discussed at the 1977 and 1978 Scarborough Conferences.

Pilot programmes in Nassau County, New York and in Maine have demonstrated beyond doubt that AFP testing in pregnancy can be carried out successfully on a regional basis within the structure of the US health care system. There is no question that such testing, properly applied, represents a major step forward in the identification and management of a variety of high-risk pregnancies, including major malformations (e.g. anencephaly, spina bifida, omphalocele, congenital nephrosis), twins, molar pregnancies, missed abortions, fetal demise, and pregnancies at high risk for premature delivery. If the full potential for this testing is to be realized, it must be actively supported and carefully executed.

Towards that end, plans are being made for a third Scarborough Conference to be held in June 1980, and emphasis is to be placed on the process of regionalisation. We hope that government agencies will have taken purposeful and positive action by that time; representatives from that sector will be invited to participate.

Yours faithfully,

JAMES E. HADDOW

EDWARD M. KLOZA

Foundation for Blood Research, Scarborough, Maine, US.

Recombinant DNA and induced tumours

SIR, — In the editorial (7 June, page 461), in which you call on the US Congress to act on recombinant DNA legislation (in itself a questionable intervention by *Nature*), you base your recommendation, in part, on risk experiments which "have indicated hitherto unanticipated potential hazards (such as the ability of DNA strands to induce tumours in laboratory animals)". Since the "risk experiments" presumably referred to were from our laboratories, we would like to comment.

As part of the polyoma risk assessment experiment (*Science* 203, 883 and 887; 1979), which was designed to monitor the possible transfer of potentially infectious DNA out of *E. coli* and into susceptible mammalian cells, we inoculated baby hamsters with polyoma virus DNA (Israel *et al.*, *J. Virol.* 29, 990; 1979) and with polyoma DNA contained in recombinant molecules (Israel *et al.*, *Science*, in the press). In brief, the two types of polyoma DNA showed comparable oncogenicity; when contained in the *E. coli* K-12 host, which was the experimental system, the recombinant DNA did not induce any tumours.

First, it seems misleading to state that "DNA strands" can induce tumours, without the qualification that you are referring to DNA from a tumour virus. Second, this finding would have been "unanticipated" only by persons who both know little or nothing of tumour virology, and who did not even read the material. Tumourigenesis and cell transformation by DNAs from various tumour viruses, including polyoma virus, have been reported many times over the past 19 years; these findings are widely known, and are well referenced in our paper in

the *Journal of Virology*, as well as in the publications cited there.

Since the example cited by *Nature* has no substance, we must wonder if there are any "unanticipated potential hazards" at all. Indeed, the 31 May issue (page 360) quotes one of us as saying "I do not know of a single piece of new data that has indicated that K-12 recombinant DNA research [meaning those experiments not prohibited by the US guidelines] could generate a biohazard." This remains our view, despite *Nature's* careless assertion to the contrary.

Major recombinant DNA policy recommendations based on uncritically accepted misstatements made by uninformed persons are only too familiar, but one would expect *Nature* to employ higher standards.

Yours faithfully,

WALLACE P. ROWE

MALCOLM A. MARTIN

National Institute of Allergy and Infectious Diseases, Maryland, US.

Preparing oral rehydration salt solutions

SIR, — Even if packets of oral rehydration salts are available (29 March, page 389), are illiterate village people capable of preparing the correct concentration of solution? If not, then they must be helped — in high concentrations the salts could be hazardous and in low concentrations they are useless. The ingredients, brown sugar, sugar, salt and sodium bicarbonate are available in most rural areas; but the 'pinch', spoon and drinking glasses are so variable that they cannot be reliably used for measuring.

The Indonesian spoon mentioned in the article is suitable for measuring the solids, but how is the fluid measured? The recommended beer bottle is not available in Muslim countries. I would, therefore, like to suggest that the UNICEF oralyte packet should contain a plastic bag, with a marked narrow neck like a bottle and instructions for measuring the fluid, and that the relevant agencies distribute cheap half-litre drinking glasses in remote villages where the UNICEF packets are not available. Three indentations should be made on the bottom of the glass: one for measuring each of sugar, salt and sodium bicarbonate. The level of the water required should be marked on the glass and instructions together with the limitations of oral rehydration should be engraved on it in local language.

The glasses should be distributed free or at a subsidised cost. They will usually be used for drinking water.

Yours faithfully,

MOSLEMUDDIN KHAN

International Centre for Diarrhoeal Disease Research, Bangladesh.

Erratum

In Norman Dombey's article (26 July, page 270) the sentence "So in a fast reactor of the design envisaged in CFR1 or Superphenix . . . there will be no overall negative Doppler coefficient" should conclude " . . . there will be an overall negative Doppler coefficient".

news and views

Hypercycles and the origin of life

from John Maynard Smith

PERHAPS the most difficult step to explain in the origin of life is that from the replication of molecules (RNA for example) in the absence of specific proteins, to the appearance of polymerases and other proteins involved in the replication of RNA and themselves coded for by that RNA. Suppose we start with a population of replicating RNA molecules. Without specific enzymes the accuracy of replication is low and hence the length of RNA which could be precisely replicated small. Before replication can be reasonably accurate, there must as a minimum be a specific polymerase, as well as synthetases and tRNAs, which in turn implies an RNA genome of considerable length. Thus, even if one supposes an initially very limited set of codons, one cannot have accurate replication without a length of RNA, say, 2,000 or more base pairs, and one cannot have that much RNA without accurate replication. This is the central problem discussed in a series of papers by Manfred Eigen and Peter Schuster (*Naturwissenschaften* 64, 541; 1977; 65, 7; 1978; 65, 341; 1978), proposing the 'hypercycle' as a necessary intermediate stage.

First, imagine a population of replicating RNA molecules, lacking genetic recombination, but with a certain 'error' or 'mutation' rate per base replication. Very roughly, if more than one mutation occurred per molecular replication, the population would come to consist of a random collection of sequences. But if less than one mutation occurred per replication, and if one sequence was 'fitter' than others in the sense of being more stable and/or more easily replicated, then a population would arise whose sequences were centred around this optimal one. Such a population Eigen and Schuster call a molecular 'quasi-species'. If the mutation rate per replication was much less than one, and if fitness differences were large, most molecules would have the optimal sequence. As the mutation rate increased, the proportion of the population with the optimal sequence would fall, until a critical mutation rate was reached, above which sequences would be random.

The replication of RNA molecules in the absence of specific replicases has not yet been achieved in the laboratory; it may

John Maynard Smith is Professor of Biology, University of Sussex.

initially have depended on the presence of nonspecific random sequence polypeptides. At this stage, the error rate would depend solely on the energy levels of base pairing between G and C, and between A and U, and would be such that the maximum length of a quasi-species would be 10–100 base pairs; even this length would require the RNA to be G–C rich. In contrast, RNA replication by the enzyme coded for by the RNA phage Q β is accurate enough to permit a genome of 1,000–10,000 base pairs, which would be sufficient to code for a primitive protein-synthesising machinery. The next stage in the evolution of accuracy of replication, achieved by prokaryotes, was the recognition and correction of mismatches in DNA replication, a process which depends on recognising which is the old strand and which the new. With such proof-correcting, a genome of at least 10⁶ base pairs, and perhaps 10⁸–10⁹ base pairs, can be maintained. It is the gap between the 10–100 bases, without specific enzymes, and the 1,000–10,000 bases, with specific enzymes, that Eigen and Schuster aim to bridge.

To understand their proposal, imagine that you wish to replicate the message GOD SAVE THE QUEEN. Counting 5 bits per letter, this is a total of 75 bits, and a maximum of 25 bits per word. Supposed the error rate was, say, 1/50 per bit. You start with a population of messages, copy each of them with that error rate, and then discard half of the population, applying a rule of selection so that messages with many errors are more likely to be discarded. Despite this 'natural selection', the population would steadily accumulate errors. The mutation rate is too high.

Suppose, therefore, that after each replication you selected word by word, discarding half the quasi-species GOD, and similarly for each other word. Since no word contains as many as 50 bits, you could now maintain a meaningful message. However, there is an important reason why this is *not* an adequate model of the replication of a set of RNA molecules. Thus, to maintain complete messages you would have to ensure in each generation that you selected equal numbers (or approximately so) of each word. But suppose the words competed, as RNA molecules would compete for substrates. Then before long only one quasi-species

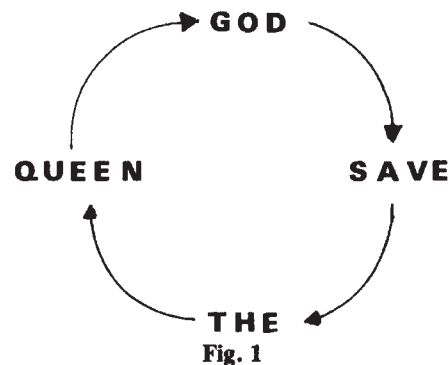


Fig. 1

Hence if the words are strung together into a single selected message (analogous to a single RNA molecule), the message is too long to replicate; if they are replicated separately, competition between the words destroys the message. You can escape from this dilemma by arranging the words in a 'hypercycle', as in Fig. 1. In this cycle, the *rate* (not the accuracy) of replication of SAVE is increased by the number of GODs, of THE by the number of SAVES, of QUEEN by the number of THEs, and, coming full circle, of GOD by the number of QUEENS. Each word, or quasi-species, is selected independently, according to its own fitness. This is a hypercycle. It can be shown that only by linking together a set of replicating quasi-species in this cyclical way can the stability of the whole be maintained.

A possible way in which a simple 'two-word' hypercycle might be realised in molecular terms is shown in Fig. 2. (Eigen and Schuster would not insist on this particular realisation; it is intended only to show the kind of thing they have in mind). It is necessary to describe a cycle in some detail, so as to be able to discuss the difficulty which arises in explaining the further evolution of such cycles. I₁ and I₂ would remain — GOD only, or THE only, if short words replicate faster.

are two RNA quasi-species. Both the + and — strands must be good replicators. RNA molecules have a phenotype, because they fold up, and folding patterns affect both survival (resistance to hydrolysis) and replication rate. Since both + and — strands must replicate, they are likely to have similar folding patterns. Further, I₁ and I₂ may be descendants of the same ancestral quasi-species, in which case they will resemble one another.

In this particular realisation, the

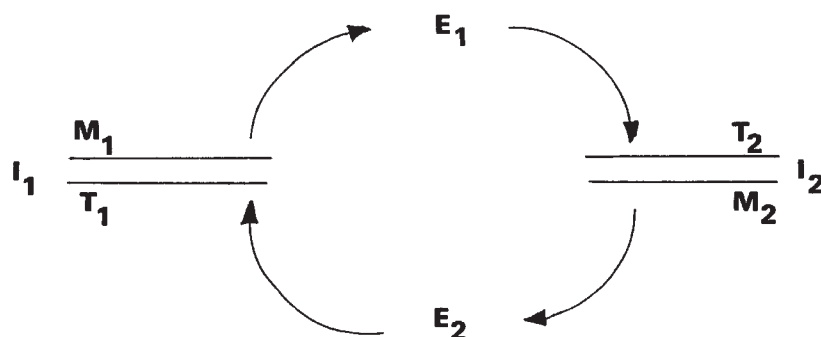


Fig. 2

— strands T_1 and T_2 are supposed to be 'adapters' or precursors of tRNA, coupling with specific amino acids, and having an anticodon loop. The + strands M_1 and M_2 are 'messengers', composed of only two kinds of codon, and coding for two proteins E_1 and E_2 which act as replicases for I_2 and I_1 respectively. Thus the model assumes that some kind of translation is possible without synthetases or ribosomes, and that the code is arbitrary from the start (that is, that there is no chemical constraint on which triplet codes for which amino acid). It is not an essential feature of the argument that messenger and transfer molecules be the + and — strands of the same quasi-species, but if they were not, then a hypercycle of more than two elements would be needed.

One difficulty in conceiving how such a cycle might arise is that it requires that each of two 'messengers' (or of more than two in the case of a longer cycle) should programme a replicase for the other: Eigen and Schuster argue that this becomes less implausible if it is remembered that I_1 and I_2 may be descended from members of the same molecular quasi-species. Initially, the members of a quasi-species would be sufficiently alike to share the same replicase. Hypercyclic organisation could then arise by the gradual differentiation of a single quasi-species into two or more.

I now digress to give the authors' views on the origin of the code. They argue that primitive RNA molecules must have been G-C rich (for reasons already mentioned). Further, in the absence of ribosomes the message must have been one which could only be read 'in frame' (Crick *et al. Origins of Life*, 7, 389; 1976). These two conditions imply that the earliest messages consisted of strings either of GNC codons or CNG codons, where N stands for any base. The authors use an argument based on wobble to prefer the former codon type. Hence, from arguments based on the stability, replication and translation of nucleic acids, they conclude that the first two codons were GGC and GCC, followed by GAC and GUC. This conclusion agrees nicely with the likely abundances of amino acids in the primitive oceans. By far the most abundant amino acids in simulated prebiotic synthesis (reviewed by Miller and Orgel, *The Origins of Life on Earth*, Prentice-Hall, 1973) are glycine and alanine (today coded by GGC and GCC),

the next commonest being aspartic acid (GAC) and valine (GUC).

Returning to the main theme, the following question arises. Given that a hypercycle ensures the accurate replication of a larger total quantity of information, how will it evolve further? Consider the particular realisation in Fig. 2. There are three kinds of mutation which might occur in I_1 (and a similar set in I_2):

(1) Mutations making I_1 better (or worse) at replicating, for example by becoming a better target for E_2 . (I_1 might also become a better target for E_1 ; if this process went too far the hypercycle would be disrupted).

(2) Mutations making E_1 a better (or worse) replicator of I_2 .

(3) Mutations making T_1 a better adapter.

If there is not compartmentalisation, only mutations of type (1) would be incorporated by selection. Each quasi-species in the hypercycle would evolve independently. There would be no selection favouring mutations of types (2) and (3), although such mutations would be needed before the speed and accuracy of replication would improve sufficiently to permit the genetic information to be united in a single 'chromosome'. There is a natural analogy here to an ecosystem. Imagine an ecosystem consisting of grass, antelopes and lions. The more grass there is, the more rapid is the multiplication of antelopes. Similarly, antelopes encourage the multiplication of lions, and (by some stretch of the imagination) lions, by fertilising the ground, encourage the growth of grass. The point of the analogy is this. Natural selection will favour mutations in grass which increase the fitness of individual grass plants, but will not favour changes in grass making it more edible to antelopes.

How then can a hypercycle evolve characteristics which favour the growth of the cycle as a whole, rather than merely its constituent parts? So long as there is no compartmentalisation, it cannot. For natural selection to act, there must be individuals. In the present context, this seems to require that the compartments of a hypercycle be enclosed in a membrane to form a proto-cell. If these proto-cells grew and divided — by some kind of budding or fission — at a rate which increased with the growth rate of the enclosed hypercycle,

then selection would favour mutations making the hypercycle more efficient. In the example of Fig. 2, mutations of types (2) and (3) would be favoured.

Clearly, these papers raise more problems than they solve. Their merit is that they put in sharp terms the problem raised by the relatively inaccurate replication of nucleic acids, and, in the hypercycle, they suggest a way in which a relatively complex structure could be replicated despite this inaccuracy. □

More twists and turns to recombination

from a Correspondent

CURRENT understanding of the molecular basis of homologous DNA recombination has evolved from analyses of the progeny of fungal and bacterial genetical crosses, the isolation of mutants deficient in some or all of the steps, electron microscopy of DNA intermediates generated *in vivo* or *in vitro* and most recently from biochemical studies of the proteins involved acting on specific substrates. The integration and transposition of IS elements, transposons and bacteriophage Mu into bacterial chromosomes has also stimulated interest in non-homologous recombination. The five EMBO recombination workshops have reflected this evolution and the latest one contained a fascinating mixture*.

Genetical analysis of fungal tetrads and octads provided support for the Meselson-Radding recombination model, which proposes initiation by invasion of a single DNA strand from one duplex into a second one, resulting in asymmetrical hybrid DNA. Symmetrical hybrid DNA can then be generated by the Siegel-Alberts isomerisation. The pattern of gene conversion and postmeiotic segregation at the *b₂* locus of *Ascombolus* is consistent with asymmetrical initiation of recombination from a preferred site outside the locus followed by a transition to a symmetrical structure, with the probability of repair of mismatched bases in the hybrid DNA decreasing with increasing distance from the initiation site (J. L. Rossignol, Université de Paris-Sud, Orsay). The positions of strand crossing-over and hybrid DNA during recombination at the *arg4* locus of yeast is, however, contrary to the model's predictions and evidence for symmetrical hybrid in yeast is lacking (S. Fogel, University of California, Berkeley). Although the direction of correction can be random in fungi, there was agreement that correction to the invading strand information is strongly preferred in yeast (Fogel; P. Hastings, University of

*The fifth EMBO Workshop on DNA Recombination, was held at Nethybridge, Scotland from 25-30 June, 1979.

Edmonton, Edmonton) and in *Sordaria*, correction may influence the likelihood of nearby cross-overs (H. Sang, Harvard University).

Repair of mismatched bases when artificial heteroduplex λ DNA molecules transfect wild-type hosts was explored (M. Radman, Université Libre de Bruxelles, Brussels). Correction occurs on the unmethylated strand, an observation of importance for correction of non-complementary bases inserted during DNA replication when parental and nascent strands are differentially methylated. During replication in *dam* mutants (adenine methylation blocked), correction can occur on both strands, perhaps leading to overlapping recombinogenic excision tracts, explaining the hyper-Rec phenotype of these mutants. Transfection of the heteroduplexes into mutator *mut* H, L and S hosts results in the loss of preferential correction. Furthermore, λ crosses under non-replicating conditions revealed that crossing-over between widely spaced markers was similar in wild-type and *mut*S hosts, whereas mismatch repair in the λ P gene was reduced 10-fold in the latter host, suggesting that mismatch repair is specifically blocked in these mutator strains.

Chi mutations in λ confer large plaque size on *red γ ⁻* mutants of λ infecting *recA* proficient hosts and also stimulate recombination in their vicinity in the λ genome. They also occur naturally in λ , *Escherichia coli* and yeast. It was suggested that the basis of their strong directionality of action might lie in the directional packaging of λ DNA and that during attempted packaging, they might resolve recombination intermediates, leading to exclusion of one of the possible products. From limited sequence data, the number of nucleotides common to all chi sequences is not yet certain. It could be as high as 23, which would not be expected to occur randomly at the frequency observed in yeast, suggesting the exciting possibility that they might be universal recombinators (F. Stahl, University of Oregon, Eugene).

The development of *in vitro* recombination systems, retarded by the lack of a versatile assay, makes the purification and characterisation of the proteins involved possible. A rapid filter-binding assay has resolved several peaks of DNA circle fusion activity after DEAE chromatography of *E. coli* extracts and one has been purified further. It is, however, present in equal amounts in both *recA⁻* and wild-type extracts and so its relationship to *recA* dependent recombination *in vivo* is obscure (D. Dressler, Harvard University). A different approach assays T7 DNA recombination *in vitro* and two pathways have been delineated. In one, purified gene 6 exonuclease catalyses recombination between two DNA duplexes or between one duplex and single strands of another. The renaturing activity of the T7 single-stranded DNA binding protein is also

required (P. Sadowski, University of Toronto).

The *recA* protein of *E. coli* is the subject of intense study because the cloning of the gene enables large amounts of the protein to be purified. It is a DNA-dependent ATPase, unwinds DNA and catalyses the insertion of homologous linear single-stranded DNA into superhelical or linear duplex DNA to form D-loops, a putative early stage in recombination (C. Radding, Yale University). On the basis of studies using ATP and a non-hydrolysable analogue, the mechanism proposed was ATP-dependent binding to homologous single strands (without ATP hydrolysis), the resulting complex unwinding the duplex DNA and annealing the single strands into the duplex, accompanied by ATP hydrolysis. Perhaps similar success will follow the cloning of the *recF* gene (A. J. Clarke, University of California, Berkeley). Recombination promoted by the λ *int* gene, which ensures the integration of λ DNA into a specific site in the host chromosome, can also be accomplished *in vitro*. The purified *int* protein has now been shown to be a topoisomerase (H. Nash, National Institutes of Health, Bethesda). It resembles eukaryotic topoisomerases and does not require attachment sites in its substrate for this activity, but its preferred binding to the phage attachment site and left attachment site may focus its activity at the specific region needed for λ integration and excision. Thus, whereas it was envisaged that recombination would require several enzymes acting sequentially, the biochemistry suggests that rather simple, but multi-functional ones can accomplish at least some initial events on their own.

The integration and transposition of IS sequences, transposons and bacteriophage Mu involves specialised or non-homologous recombination, and transposition is non-reciprocal because they are inserted into a recipient without their loss from the donor. Thus DNA synthesis is an integral part of transposition. Partial deletions in Tn3 were used to analyse transposon functions required for transposition (D. Sherratt, University of Sussex). Some deletions blocked transposition at an intermediate stage where a Tn3-containing plasmid was physically linked to a recipient plasmid. In some cases such intermediates were resolved by the host *recA* system which was, however, less efficient than the transposon's own system. A sequence within 250 base pairs of a gene coding for a 19,000 molecular weight protein was also mapped which is required in *cis* to resolve cointegrates or Tn3 dimers in a plasmid. A search for *E. coli* mutants reducing IS1-mediated deletion formation in the *gal* operon revealed a trans-acting locus named *del* (P. Nevers, University of Freiburg). From further studies, it was concluded that *del⁺* might be a natural IS1 sequence linked to a strong promoter acting in

trans to boost IS1 functions at a second insertion site in the *gal* operon.

Bacteriophage Mu shares many properties with transposons, such as a requirement for replication to transfer Mu from one site to another (A. Bukhari, Cold Spring Harbor Laboratory, New York). Also Mu can transpose from one plasmid to another equally efficiently in *recA⁺* and *recA⁻* hosts, but in the latter, cointegrates are detected which can be resolved by the *recA⁺* system. Two Mu deletion mutants also form cointegrates, but the one with deleted early genes is unable to transpose into the host chromosome (D. Leach, University of Sussex). When Mu is induced, the *G* region of about 50% of the progeny is inverted, preventing growth of Mu on *E. coli*. This puzzle has been resolved because this loss of infectivity is matched by a gain in ability to infect other bacteria. *G* inversion apparently activates different *S* genes, coding for host specific phage tail fibres, by way of a promoter at the left of *G*. Inversion also increased read-through of a flanking gene *mom* which antagonises host-specific modification of Mu DNA so that phage with a new host range are less likely to have their DNA restricted.

Transposon-like sequences may also exist in yeast. The *his4-912* mutation maps outside structural genes, is polar but is not a nonsense, frameshift or deletion and revertants are unstable and cause chromosomal rearrangements. Restriction fragments containing the region from *his4-912*, a revertant and wild-type decrease in length in that order, the revertant being longer than the wild-type by 250 base pairs, a sequence found in many yeast restriction fragments. It was suggested that *his4-912* contains a transposable element in the promoter which is excised in the revertant leaving a 250 base pair terminal repeat (G. Roeder, University of Toronto). The 'Cassette Model' for control of the mating type in yeast is also consistent with a transposon-like mechanism. Cloning of the α and A mating type alleles indicates that the former is 200 base pairs longer, but the degree of internal homology between them remains to be determined (A. Klar, Cold Spring Harbor Laboratory).

The possible involvement of recombination in carcinogenesis was proposed by Radman and A. Kinsella. TPA, a strong promoter of carcinogenesis, causes a large increase in sister chromatid exchanges in cultured CHO cells. They propose that carcinogenesis is initiated by mutations which remain silent because of the presence of intact DNA on the homologous chromosome. If recombination occurs, perhaps stimulated by TPA, there is an opportunity for the loss of the wild-type gene, permitting expression of the carcinogenic mutation. Their proposal was strengthened by the finding that the anti-tumour agent antipain, prevents sister chromatid exchange. □

Structure and control of phosphorylase

from C.C.F. Blake

It is perhaps not really surprising that the promise that X-ray crystallography can make a vital contribution to the understanding of enzyme regulation, by being able to define the various structural states through which the activities of control enzymes are regulated, is taking a long time to be completely fulfilled. The enzymes by their very nature are large, their conformational states can be various, and can differ from one another to such an extent that a number of crystal forms may have to be analysed separately (if they can be obtained at all). Of the four regulatory enzymes whose X-ray study is well advanced: glycogen phosphorylase, hexokinase, phosphofructokinase and glyceraldehyde phosphate dehydrogenase (all members of the glycolytic pathway), phosphorylase is under the most elaborate control. This suggests that the analysis of its structural and regulatory properties now under way by Fletterick and Madsen in Edmonton, and Johnson in Oxford will be especially valuable in providing information on the regulation of metabolism.

Glycogen phosphorylase is, of course, the classic example of hormonal control of a metabolic enzyme through the mediation of the adenylyl cyclase system. Phosphorylase *b* is converted to phosphorylase *a* by the phosphorylation of Ser 14 by a specific kinase; a reaction that is reversed by a specific phosphatase. This hormonal regulation is linked to further control by the levels of a number of metabolites, including AMP, ATP and glucose-6-phosphate. Phosphorylase *b* requires AMP for activity and is allosterically inhibited by ATP and glucose-6-phosphate. In resting muscle the levels of these metabolites are such that this form of the enzyme is inactive. The conversion of the *b*-form to the *a*-form results in an enzyme that no longer requires AMP for activity and is not subject to allosteric inhibition. However, it is strongly inhibited by glucose, for which it is regarded as a receptor in the liver, and which promotes the conversion of the *a*-form to the *b*-form by the phosphatase. In addition phosphorylase is in intimate physical association with the glycogen particles, and this association also seems to affect the enzymic activity.

X-ray studies are being carried out on crystals of phosphorylase *b* grown in the presence of its activator IMP at 3 Å resolution (Johnson *et al. Proc. 11th FEBS Mtg.* 42, 185; 1978; Weber *et al. Nature*, 274, 288; 1978) and of phosphorylase *a* in the presence of its inhibitor glucose at 3 Å resolution (Fletterick *et al. J. biol. Chem.* 251, 6142; 1976) which have now been extended to 2.5 Å resolution (Sprang &

Fletterick, *J. molec. Biol.* 131, 523; 1979). It was extremely valuable that the complete sequence of the 841 residues of rabbit muscle phosphorylase was determined by Neurath's group (Titani *et al. Proc. natn. Acad. Sci. U.S.A.* 74, 4762; 1977) at the time when the electron density maps were being interpreted. The 97,400 molecular weight subunit of phosphorylase has its polypeptide chain organised largely into alternating α - and β -structures, like all the other glycolytic enzymes whose structures are known. The folding of the polypeptide chain divides the large subunit into three domains: an N-terminal domain of about 320 residues, a central domain of 160 residues, and a C-terminal domain of about 360 residues. These three domains are located at the apices of a roughly equilateral triangle giving the subunit a compact shape. The interface of the functional dimers in the crystals is formed by the N-terminal domains and includes two large projections, known as the 'tower' and the 'cap' that extend beyond the molecular two-fold symmetry axis to make extensive contacts with the other subunit. A most striking feature of the subunit is the location of the active site at the bottom of a deep, narrow cavity near the centre of the subunit at the point where the three domains come together. The pyridoxal phosphate, whose presence was so long a mystery, is bound as a Schiff's base to Lys 679 close to the active site in apparent confirmation of proposals that it is directly involved in the catalytic process. In phosphorylase *a*, the glucose inhibitor is bound in the active site.

Rather surprisingly the crystals of phosphorylase *a* and *b* are closely isomorphous, suggesting no large structural difference between them. Unfortunately no detailed comparisons of the two structures have yet been made, but it seems that a major difference lies in the N-terminal region which in the *b*-form is disordered, but becomes ordered in the *a*-form because the phosphate group attached to Ser 14 makes an ionic interaction with Arg 69 located on the N-terminal helix. Soaking experiments with substrates, activators and inhibitors have revealed further binding sites (Kasvisky *et al. J. biol. Chem.* 253, 1290, 3343; 1978; Johnson *et al.* unpublished).

The allosteric activator site lies on the N-terminal domain at the subunit interface. AMP and ATP located in this site make contact with residues from both subunits. Maltotriose and maltoheptose bind at a site on the central domain, which has consequently been termed the glycogen storage domain, but do not seem to bind at the active site. Finally, AMP, ATP and the inhibitors caffeine and inosine, bind near the active site in a position that probably marks the allosteric inhibitor site.

Despite the success in locating these five ligand binding sites, the only evidence for associated conformational change has been obtained from crystals of phosphorylase *a* in the presence of the virtual substrate, 5-thio- α -D-glucopyranosyl phosphate, and maltoheptose (Madsen *et al. J. biol. Chem.* 253, 9097; 1978). To prevent the disintegration of the crystals that usually accompanies attempts to bind substrate, they were lightly cross-linked with glutaraldehyde and allowed to anneal after the addition of the substrate analogue. Difference maps at low resolution revealed that the segments in the N-terminal half of the molecule had undergone conformational change, possibly representative of the R $\rightarrow$ T transition. The most intriguing aspect of these conformational changes is that they appear to link together the five major ligand binding sites. These structural changes seem to bear obvious relationships to the known homotropic and heterotropic cooperativities of the enzyme.

The work on phosphorylase is an exemplar of X-ray studies on enzymes, regulatory or not, both in its successes and difficulties. As is usual the structural results are clear, and revealing. The division of the molecule into three domains of specialised function is characteristic, as is the location of the active site at the triple domain interface. The domain/subunit geometry lays the groundplan for regulation by defining, broadly, the structural pathways of communication between the various control sites and the active sites. However, as is also usual, the definition of function, that is how the communication actually works, seems more problematic even though in this case some low resolution results have been obtained with phosphorylase *a*. This is because the structural changes induced by ligand binding may be partly, or wholly, inhibited by the crystal lattice. Ideally separate structural analyses of crystals representing each stable structural state are required, but that may not be possible because certain states may not crystallise. Function is always the most difficult aspect of X-ray studies, and one that requires much further effort. □

Erratum

In the article 'The Oceanography of Fjords' (*News and Views* 280, 273; 1979) the figure legend is incorrect. The upper figure represents the laboratory simulation and the lower the tidally-generated flow in a fjord, and not as stated.

C.C.F. Blake is in the Laboratory of Molecular Biophysics, Department of Zoology, University of Oxford.

Theoretical difficulties with the binary pulsar?

from Malcolm MacCallum

THE agreement of the rate of change of period of the binary pulsar with the usual quadrupole formula for gravitational radiation, reported by J. H. Taylor, L. A. Fowler, and P. M. McCulloch (*Nature* 277, 437; 1979), has been hailed as a triumph for general relativity. Unfortunately it has been known to specialists for some time that the theoretical predictions are themselves open to question.

J. Ehlers, A. Rosenblum, J. N. Goldberg and P. Havas drew attention to the unsatisfactory nature of all existing derivations of the formulae in an article in the *Astrophysical Journal* in 1976 (208, L77). Concerning these arguments, Taylor *et al.* remark that "our data suggest that any such inaccuracy is not very large", but there is of course the more intriguing possibility that the observations disagree with the correct general-relativistic formula. Indeed, what some specialists fear most is that the usual formula will prove correct to better than 10%, with the consequence that any future objections to non-rigorous arguments in relativistic astrophysics and cosmology will be dismissed as merely pedantic quibbles.

In deriving the standard formula, L. D. Landau and E. M. Lifschitz (in *The Classical Theory of Fields* third edition, Pergamon, 1971) wrote "In principle, all the calculations are completely analogous to those which we carried out for electromagnetic waves." It is this analogy which has given the quadrupole formula its strong intuitive appeal to physicists, but it also emphasises the underlying reasons for some of the experts' doubts, namely that it uses an essentially linear approximation in a flat space-time background when the full theory is non-linear and uses curved spaces.

Since 1976 renewed efforts have been made to reach an agreed alternative to (or confirmation of ?) the usual formula, based on more sophisticated approximations, and the Gregynog meeting* provided an opportunity to review progress. Some of the most illuminating results reported arose from model calculations using scalar waves in one dimension (whose relevance is accentuated by a recent preprint from J. M. Stewart (University of Cambridge) showing that perturbations of algebraically special space-times — which include most of the physically interesting cases — can be represented in terms of scalar potentials).

At great distances from the sources, the field is usually assumed to propagate along

flat space light rays and to be expandable in inverse powers of the radial distance r . Stewart, speaking about definitions of mass at infinity, reminded the audience that, however small the mass involved, light rays in the field of a spherical body (a Schwarzschild metric) would diverge logarithmically from those of the corresponding flat space. M. Walker (Max-Planck-Institute, Munich) reported an example in which logarithmic terms appear in the field in the far past.

One way of relating the field to the sources is through a retarded Green's function integral. In relativity, however, such retarded solutions are not simply related to natural choices of boundary condition, such as 'no incoming radiation in the far past' or 'only outgoing radiation in the far future'. Walker gave an example in which the retarded solution involved incoming radiation in the far past. B. Schmidt (Max-Planck-Institute, Munich) presented a string and spring example in which the amount of incoming radiation affected the damping rate, and speculated, amusingly, that the binary pulsar measurements might mean only that if general relativity is true, the pulsar is receiving no incoming radiation. Schutz proposed avoiding the boundary conditions by assuming that at an initial instant the field external to the system, being unknown, should be taken to be zero — the average of all possible values.

The dynamics of the source has to take account of the reaction from the field. Schutz showed that in his approach the usual radiation reaction expression acquired an extra term, which he related to the need to compensate for coordinate choice effects. More detailed examinations of the internal dynamics of the sources usually follow one of two approximation schemes.

The first of these is the 'fast motion' approximation. D. Christodoulou (Max-Planck-Institute, Munich) reported on some rigorous estimates of errors in this scheme. A. Rosenblum (Temple University, Philadelphia) reported that the well-known discrepant result of Havas and Goldberg (that two bodies in a binary system spiral away from, not towards, each other) was due to neglect of non-linear terms comparable with those retained, and presented his own scattering calculations, which gave about double the result predicted by the quadrupole formula. He hopes to extend these calculations to bound systems.

The second method is the 'slow motion' technique, which involves expansion in powers of a small parameter ϵ , essentially the ratio of a characteristic velocity in the system to the velocity of light. The calcul-

ations introduce terms in ϵr , and are therefore valid only in a limited region. The boundary conditions for the slow motion approximation should be supplied by an outer, wave zone, solution, which itself must use the light rays of the curved space, and not a flat space approximation. Both Dixon and J. L. Anderson (Stevens Institute, Hoboken, New Jersey) have carried out model calculations of this type, joining the inner and outer regions by the method of matched asymptotic expansions. This method was introduced to relativity by W. J. Burke (*J. Math. Phys.* 12, 401; 1971), but has been familiar in other contexts for some time, and is even known to be rigorous in certain limited circumstances. Because their calculations differed in detail, Dixon and Anderson, although in broad agreement, emphasised different aspects of the results.

Dixon's model calculation showed, contrary to some expectations, that slow motion expansions can include effects arising from tail terms of a retarded Green's function (crudely, effects propagating at a speed less than that of light), and that radiative effects can appear in even as well as odd orders in the expansion. Both authors found that the matching introduced terms in the inner region logarithmic in ϵ . Anderson noted that approximations good for one feature may be bad for another (in his example, the damping and phase respectively), though I understand that at the subsequent Dublin meeting ('Current Problems in General Relativity', 2-6 July, 1979) P. Parker (Syracuse University) suggested that this could be overcome using results by Hörmander (*Acta Math.* 127, 79; 1971).

The problem of what to approximate recurred in a final discussion excellently led by Stewart. It was agreed that intermediate concepts like energy (awkward to define in the full theory) and radiation reaction force should be avoided if possible, and attention focused on directly measurable quantities (like the binary pulsar period changes). Care would be needed to avoid tailoring approximations to produce the 'right' results, so that as many observable effects as possible should be included.

Stewart concluded the meeting by announcing a competition for a model problem, with an unambiguous numerical solution, which could act as a testbed for the different approximation schemes. It should have as many features of the full problem as possible, for example, non-linear waves, retarded coupling to a source oscillator, and back-reaction on the source. Perhaps when such a problem has been found, and solved, we will know how to tackle the binary pulsar question. Then we can decide whether the observed behaviour presents difficulties for the theory, or whether the present doubts will prove only 'theoretical' in the worse sense.

Malcolm MacCallum is a lecturer in the Department of Applied Mathematics, Queen Mary College, London.

*The third Gregynog Relativity Workshop was held at Gregynog, Newtown, Wales, on 25-28 June, 1979, under the title 'Gravitational Radiation Theory'. The organisers were B. Schutz (University College, Cardiff), G. Dixon (Churchill College, Cambridge) and M. Walker (Max-Planck-Institute, Munich).

Minding the psi's and q's

from Roger Cashmore

THE European Physical Society's International Conference on High Energy Physics* was held in Geneva this year — the 25th anniversary of CERN. Much of the data presented originated from accelerators within Europe, from the CERN SPS and ISR together with a substantial contribution from the new e^+e^- storage ring PETRA at DESY.

Much of the conference was devoted to reports on the physics of both the new quarks, (charm and bottom — c, b) and the old quarks (up, down and strange — u, d, s) as well as searches for the expected quark (top — t). First the charm quark. Within the charmonium model of the $c\bar{c}$ meson states, the claimed pseudoscalar states (1S_0 states), the χ (~ 2830) and χ (~ 3450) have long been a problem. They possessed masses, widths and M1 electromagnetic transition rates in conflict with predictions from charmonium. In the crystal ball experiment at SPEAR, an experiment particularly designed to study the electromagnetic transitions from the ψ and ψ mesons using high resolution NaI crystals, no evidence has been found for these states at the claimed mass values, although all of the other intermediate states have been observed (the $^3P_{0,1,2}$ states — the $\chi(3400)$, $\chi(3510)$ and $\chi(3550)$). Thus the embarrassing problem of having states of the grossly wrong properties has been removed but replaced by the lesser problem of the whereabouts of these mesons, vital in the $c\bar{c}$ picture. In a separate experiment at SPEAR, using the Mark II detector, data has been taken on the D^0 and $\bar{D}^0$ mesons at the ψ'' (3772), a copious source of these particles. With the high statistics obtained the first observations of the Cabbibo suppressed $D \rightarrow \pi\pi$ and KK decays have been made which are consistent with 6 quark extensions of the GIM model of incorporating new quarks into the weak and electromagnetic gauge theories.

Until this conference evidence for charmed baryons has been scant, the only observations being in neutrino experiments and one photoproduction experiment. Three experiments from the ISR reported the observation of narrow hadronic states consistent with the earlier observations. These states have appreciable diffractive production cross sections ($\sigma_B \sim 1\mu\text{b}$, where B is the decay mode branching fraction) while the more prominent decay mode of the lightest of these baryons, the Λ_c^+ (~ 2255) appears to be to $K^-p\pi^+$ although the $\Lambda_c^0\pi^+\pi^-\pi^0$ decay is observed. This conclusion is supported by evidence from SPEAR that the $p + \bar{p}$ yield increases more dramatically than the $\Lambda + \bar{\Lambda}$ yield in the vicinity of 5.0 GeV, the region where charmed baryon production is expected to occur.

Probably the most exciting prospect at the conference was the suggestion of a meson containing a bottom (b) quark. The b quark is expected to decay preferentially to a c quark and it has been surmised that this c quark might appear a fraction of the time within a ψ meson together with at least one 'strange' particle. A group (Indiana/Imperial College/Saclay/Southampton collaboration) studying ψ production in ~ 150 GeV $\pi^+\pi^-$ collisions at the CERN SPS looked at the $\psi K\pi$ final states and discovered an excess of events (~ 4 standard deviations) concentrated in a single 40 MeV bin at a mass of ~ 5.3 GeV. From a knowledge of the Υ state, assumed to be a $b\bar{b}$ meson, the mass of a meson containing a bottom quark can be estimated and falls in this vicinity. Clearly further experimental results are eagerly awaited in the hope of confirming this exciting possibility.

Finally, an expected new quark, the t quark, has not yet been revealed in e^+e^- annihilation experiments at PETRA. Thus the mass of this quark must be greater than 13.5 GeV and means that the new accelerator must be stretched to its present limit of 16 GeV in the hope of discovering it.

Nearly all known mesons and baryons can be accommodated in models of hadrons in which mesons are $q\bar{q}$ states and baryons are qqq states, leading to the puzzle of the absence of resonances composed of greater numbers of quarks, for example $qqqq$ or $qqqqq$. A few narrow structures observed in pp annihilations had been the best candidates for such states but unfortunately in a recent pp experiment at Brookhaven National Laboratory no evidence was found for the existence of one of the best of these candidates, the $S(1935)$. Thus the puzzle remains as one of the most important questions of hadron physics.

The problem of permanently confining quarks within hadrons has apparently yet to be solved despite the tremendous amount of theoretical effort directed towards QCD (quantum chromodynamics — the theory of coloured quark and gluon interactions), a possible theory of the strong interaction. The experimental evidence for jets of particles, the remains of original quarks, was beautifully demonstrated in the new results from PETRA. However the evidence for the gluons which are essential in this picture remains tentative. Data presented on the decay of the Υ (9.45) is clearly best fitted with a model which describes its decay by way of three gluons, while in the high energy e^+e^- annihilation the evidence for a third jet (a gluon bremsstrahlung from one of the quarks) was tantalising. The existence of this effect together with a quantitative confirmation of the QCD predictions will only be obtained with the acquisition of much more data during the coming year. Results were also presented on the variation of the moments of the structure functions in deep inelastic νN and

Naming names

IN a paper published in this issue of *Nature* (page 468) V. Bennett and P.J. Stenbuck describe how a red cell membrane protein is linked to the cytoskeletal spectrin through the membrane Band 2.1 protein. Bennett and Stenbuck have recently rechristened Band 2.1 protein ankyrin (Greek, anchor). But beware. The self-same protein has, by others, been rechristened nexin (Latin, to bind) and syndein (Greek, to bind). Perplexingly, the work that led to all three christenings was initiated in one laboratory, that of Professor D. Branton at Harvard.

This etymological discordance is hardly edifying. Perhaps the younger generation should devote less time to scouring dictionaries and reflect instead on other traditions. After all there was a time when the leader of an expedition could expect to have his (or even her) name enshrined in any new species discovered by juniors.

So, by way of revitalising an old tradition, I hereby propose that the names ankyrin, nexin and syndein be cast aside in favour of brantonin (from, if you must, the words branch — a subdivision — and ton — a large quantity). International Commissions on Nomenclature please note. P.N.

μN scattering, the 'classic' testing ground of QCD to date. Once again these were consistent with the predictions of QCD but fall short of a definitive proof, since other explanations are possible.

Results from all the high energy physics experiments studying the weak neutral current are now consistent with a value of $\sin^2\theta_w = 0.23$, the one parameter in the Weinberg-Salam model for a gauge theory of the weak and electromagnetic interactions. The only area in which there is any conflict with this value lies in atomic physics experiments studying parity violation in thallium and bismuth atoms. However here there are direct conflicts between the various experiments and a further sequence of experiments is clearly required before this situation is resolved.

There remains the unification of the weak, electromagnetic and strong interactions. The currently most popular models imply a 'desert' from the mass region of the W and Z intermediate bosons (~ 100 GeV) to the mass at which the various interactions have a similar strength — the grand unification mass ($\sim 10^{15}$ GeV). However A. Salam, in his summary, pointed out that this 'desert' was model-dependent and indeed other models would lead to a proliferation of particles in this region — only experiment will reveal the correct path. This represented a clear challenge to the experimental high energy physicists to understand better the properties of the particles of today and furthermore to achieve these higher energies as soon as possible with the accelerators of the future, the pp collider at CERN and, one hopes, LEP. □

Roger Cashmore is in the Department of Nuclear Physics, University of Oxford.

*Held on 27 June-4 July.

review articles

Evidence for the quantum nature of light

D. F. Walls

Department of Physics, University of Waikato, Hamilton, New Zealand

A unique property predicted by the quantum theory of light is the phenomenon of photon antibunching. Recent theoretical predictions and experimental observations of photon antibunching in resonance fluorescence from a two-level atom are reviewed.

THE advent of photon correlation experiments, as pioneered by Hanbury-Brown and Twiss¹, began a new era in the field of optics. The experiments of classical optics involved a measurement of the first-order correlation function of the electromagnetic field and as such involved only the interference between the probability amplitudes of a single photon. Phenomena such as diffraction, Young's interference experiment and spectral measurements may be categorised as being in the domain of one photon or linear optics. Photon correlation experiments represent a fundamental deviation from these experiments as they involve the interference between different photons and as such are in the realm of nonlinear optics. The quantum-mechanical interpretation of the Hanbury-Brown Twiss effect was given by Glauber². However, adequate explanations of the Hanbury-Brown Twiss effect and related photon correlation experiments have been given which invoke a classical description of a fluctuating electromagnetic field. However, Glauber² pointed out that photon correlation experiments offer the possibility of observing a uniquely quantum-mechanical effect—namely photon antibunching. We describe here the recent prediction by Carmichael and Walls³ and observation by Kimble, Dagenais and Mandel⁴⁻⁶ and G. Leuchs, M. Rateike and H. Walther (personal communication) of this unique quantum-mechanical effect in photon correlation experiments of fluorescent light from a two-level atom. This marks a dramatic change from all previous photon correlation experiments since the photon antibunching observed cannot be explained on the basis of a classical description of the radiation field. We begin with a brief outline of events leading up to the present investigations.

The quantum theory of light

The quantum theory of light beginning with Planck⁷ and Einstein⁸ played a central part in the development of quantum theory during this century. As the quantum theory was developed a sophisticated theory for the interaction of photons and electrons namely quantum electrodynamics evolved. Amongst the predictions of quantum electrodynamics was that the emission of light from an atom would experience a small shift away from the resonance line of the atom. The experimental observation of this shift, known as the Lamb shift, came as a major triumph for the quantum theory of light⁹. However, for many experiments in optics especially in the field of physical

optics the classical picture of light as propagating waves of electromagnetic radiation was adequate to describe the observations. The idea developed that it was not necessary to quantise the light field but only necessary to quantise the atoms and thus describe their interaction by a semiclassical theory. Modifications of this concept such as neoclassical radiation theory also developed. These theories can explain phenomena such as the photoelectric effect, spontaneous emission and even give estimates of the Lamb shift. (A discussion on the present status of such theories as an alternative to quantum electrodynamics is given in ref. 10.) An early attempt to find quantum effects in a physical optics experiment was made by Taylor¹¹ who performed Young's interference experiment at very low intensities such that on the average only one photon was incident on the screen at a time. Integrating over a long detection time Taylor obtained an interference pattern as predicted by classical wave theory with no evidence of any unique quantum effects. Taylor's experiment may also be explained on the basis of the quantum theory which interprets it as the interference of the quantum-mechanical probability amplitudes for the photon to go through one slit or the other^{12,13}. Essentially the experiments of physical optics were in the regime of one photon or linear optics. We need to go outside the experiments of one photon optics in order to detect any uniquely quantum-mechanical effects.

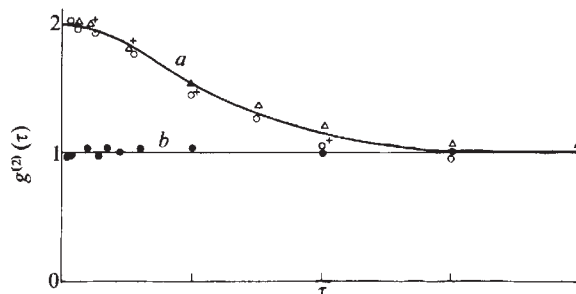


Fig. 1 Second order correlation function $g^{(2)}(\tau)$ for: *a*, thermal light; *b*, coherent light. Experimental points from Arrecchi *et al.*¹⁷.

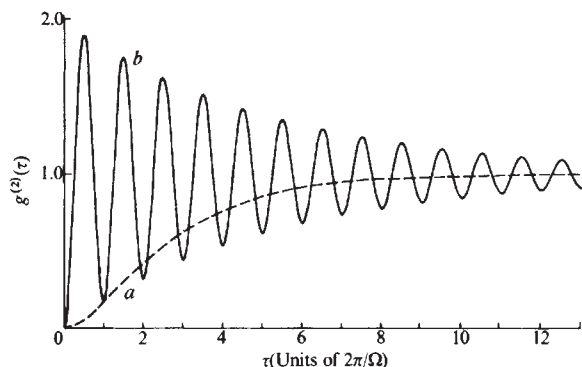


Fig. 2 Second order correlation function $g^{(2)}(\tau)$ for fluorescent light from a two level atom: *a*, low driving field intensity $\Omega \ll \gamma$; *b*, high driving field intensity $\Omega \gg \gamma$. Theoretical predictions from Carmichael and Walls³.

Photon correlation experiments

The first experiment outside the domain of one photon optics was performed by Hanbury-Brown and Twiss in 1956. They performed an intensity correlation experiment, that is a measurement of $\langle I(t)I(t+\tau) \rangle$. Although the original experiment involved the analogue correlation of photocurrents, later experiments used photon counters and digital correlators¹⁴⁻¹⁸. In essence these experiments measure the joint photocount probability of detecting the arrival of a photon at time t and another photon at time $t+\tau$. The measurements made in this experiment may be described in terms of the second-order correlation function of the radiation field introduced by Glauber² in his formulation of optical coherence theory. This is defined as

$$g^{(2)}(\tau) = \frac{\langle E^{(-)}(t)E^{(-)}(t+\tau)E^{(+)}(t)E^{(+)}(t+\tau) \rangle}{\langle E^{(-)}(t)E^{(+)}(t) \rangle^2} \quad (1)$$

where $E^{(+)}(t)$ and $E^{(-)}(t)$ are the positive and negative frequency components of the electromagnetic field respectively.

The result of a photon correlation measurement of $g^{(2)}(\tau)$ for a thermal light source using photomultipliers and digital correlators is shown in Fig. 1 curve *a*. We see that the joint counting rate for zero time delay is twice the coincidence rate for large time delays τ . That is, there is a tendency for the photons to arrive in pairs, or a photon-bunching effect. The decay time of the correlations is given by the inverse bandwidth of the light source.

If instead of a thermal light source the experiment is performed with a highly stabilised laser source one obtains a constant correlation function as shown in Fig. 1 curve *b*. This result holds even if the laser and thermal light source have the same bandwidth. Hence there is some fundamental difference between a laser and a thermal light source which may not be apparent in the first-order correlation function but which is manifest in the second-order correlation function. This effect may be understood by considering the photon statistics of different light fields.

Quantum theory of photon correlations

We shall attempt in this section to give a quantum theoretic interpretation of photon correlation experiments. For simplicity we shall concentrate on an interpretation of $g^{(2)}(0)$ which enables us to restrict our attention to a single mode field. For a single mode field we may write $g^{(2)}(0)$ in the form

$$g^{(2)}(0) = 1 + \frac{\sigma^2 - \bar{n}}{\bar{n}^2} \quad (2)$$

where $\bar{n}$ and σ^2 are the mean and variance respectively of the photon number distribution. Thermal light has a power law photon number distribution with variance $\sigma^2 = \bar{n}^2 + \bar{n}$ which gives $g^{(2)}(0) = 2$. Coherent light on the other hand as produced by a highly stabilised laser has a poissonian photon number distribution¹⁹⁻²¹ with $\sigma^2 = \bar{n}$ leading to $g^{(2)}(0) = 1$. This clearly yields the results shown in Fig. 1 for $g^{(2)}(0)$. To include the decay time of the correlation function one must consider many modes. For a field which has a photon number distribution narrower than a Poisson ($\sigma^2 < \bar{n}$) it is possible in principle to obtain a $g^{(2)}(0) < 1$. That is the photon coincidence rate for zero time delay is less than that for large time delays τ . This is the opposite effect to the photon bunching observed for thermal light and has been called photon antibunching. This represents a reduction in the photon number fluctuations below that of a poissonian distribution. The extreme case of photon antibunching would be a light beam where the photons arrive evenly spaced. For a single mode field this is equivalent to having a fixed photon number N which gives $g^{(2)}(0) = 1 - 1/N$.

Let us now consider a classical description of the electromagnetic field. We consider a field described by a fluctuating amplitude E . These fluctuations are taken into account by introducing a probability distribution $P(E)$ for the complex field amplitude. A calculation of $g^{(2)}(0)$ yields

$$g^{(2)}(0) - 1 = \frac{\int P(E)(|E|^2 - \langle |E|^2 \rangle)^2 d^2E}{\langle |E|^2 \rangle^2} \quad (3)$$

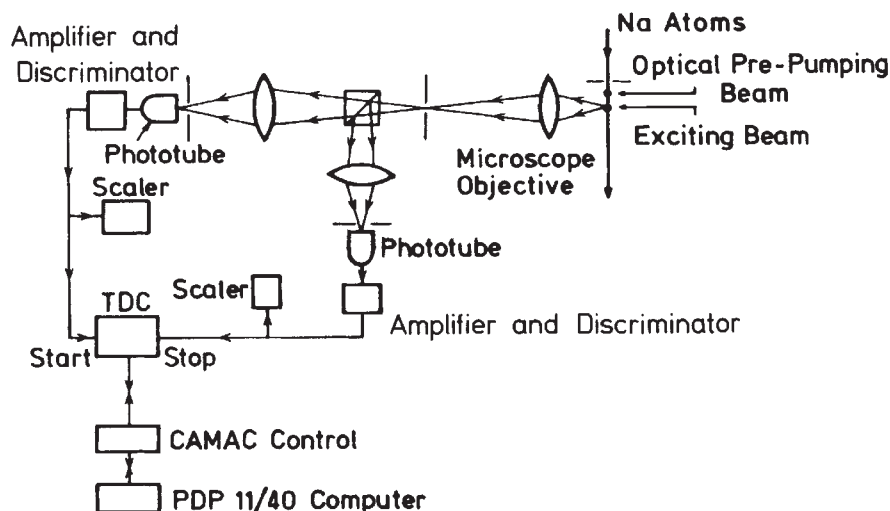


Fig. 3 Experimental arrangement used by Kimble *et al.*⁴.

For a thermal field which has a gaussian distribution for $P(E)$ we recover $g^{(2)}(0) = 2$ whereas for a coherent field with a stabilised amplitude $P(E)$ is a delta function and hence $g^{(2)}(0) = 1$. Hence the Hanbury-Brown and Twiss effect for chaotic fields and indeed for the coherent laser field is adequately described by classical theory. However, since the right hand side of equation (3) is a positive semi-definite quantity, we see that classical theory requires $g^{(2)}(0) \geq 1$ and hence does not allow photon antibunching.

The formulation of the quantum theory closest in formal appearance to the classical theory involves the diagonal expansion of the density operator of the radiation field in terms of coherent states^{2,22},

$$\rho = \int P(\mathcal{E}) |\mathcal{E}\rangle \langle \mathcal{E}| d^2\mathcal{E} \quad (4)$$

where the coherent states $|\mathcal{E}\rangle$ are eigenstates of the positive frequency part of the electromagnetic field

$$E^{(+)}|\mathcal{E}\rangle = \mathcal{E}|\mathcal{E}\rangle \quad (5)$$

with eigenvalue $\mathcal{E}$. The function $P(\mathcal{E})$ does not have the character of a probability distribution since for certain nonclassical states of the radiation field it may be negative or highly singular.

A calculation of $g^{(2)}(0)$ using the P representation yields

$$g^{(2)}(0) - 1 = \frac{\int P(\mathcal{E}) (|\mathcal{E}|^2 - \langle |\mathcal{E}|^2 \rangle)^2 d^2\mathcal{E}}{\langle |\mathcal{E}|^2 \rangle^2} \quad (6)$$

This appears similar in form to the classical expression of equation (3), however, since $P(\mathcal{E})$ may for certain fields which have no classical description take on negative values then $g^{(2)}(0)$ may be less than unity in quantum theory.

While this particular feature of quantum theory was recognised it remained to find a light field which exhibited the property of photon antibunching. It was suggested that certain processes of nonlinear optics for example sub-second harmonic generation²³⁻²⁶ and two-photon absorption²⁷⁻³³ would exhibit photon antibunching. To date there has been no experimental verification of these predictions. It was resonance fluorescence from a two-level atom that led to the first experimental observation of photon antibunching.

Resonance fluorescence from a two-level atom

The topic of resonance fluorescence from a two-level atom has been the subject of considerable theoretical and experimental investigation. For weak incident fields the light is coherently scattered, whereas for strong incident fields when the Rabi frequency Ω ($\Omega = 2\kappa\mathcal{E}/\hbar$ where $\mathcal{E}$ is the driving field amplitude and κ is the atomic dipole matrix element) exceeds the Einstein A coefficient γ of the atom the spectrum of the scattered light splits into three peaks with the two sidebands displaced from the central peak by the Rabi frequency. This spectrum first predicted by Mollow³⁴ was verified in detail by the experiments of Wu *et al.*³⁵ and Hartig *et al.*³⁶ following an earlier experiment of Schuda *et al.*³⁷. These experiments used an atomic beam of sodium atoms which were optically pumped to prepare a pure two-level system. These atoms were then subjected to irradiation from a highly stabilised dye laser tuned to resonance with the atomic transition. This system was to prove of further interest to physicists.

Carmichael and Walls³ predicted that the second-order correlation function of the light emitted by a single atom undergoing resonance fluorescence would exhibit the property of photon antibunching. This was confirmed in subsequent calculations by Cohen-Tannoudji³⁸ and Kimble and Mandel³⁹. The results obtained by Carmichael and Walls for the second order correlation function $g^{(2)}(\tau)$ in the steady state are

$$g^{(2)}(\tau) = \left[1 - \exp\left(-\frac{\gamma\tau}{2}\right) \right]^2 \quad \Omega \ll \gamma \quad (7)$$

$$g^{(2)}(\tau) = \left[1 - \exp\left(-\frac{3\gamma}{4}\tau\right) \cos \Omega\tau \right] \quad \Omega \gg \gamma \quad (8)$$

in the limiting cases of very weak and very strong driving fields. The behaviour of $g^{(2)}(\tau)$ is plotted in Fig. 2. This system clearly displays the property of photon antibunching since in both limits $g^{(2)}(\tau)$ starts at zero. For low-intensity driving fields $g^{(2)}(\tau)$ rises monotonically to a background value of unity, whereas for high driving field intensities $g^{(2)}(\tau)$ rises above unity before reaching a steady state value of unity by damped oscillations.

This behaviour can be understood as follows. A measurement of $g^{(2)}(\tau)$ records the joint probability for the arrival of a photon at time $t = 0$ and the arrival of a photon at time $t = \tau$. Consider now the driven two-level atom as our source of photons. The detection of a fluorescent photon prepares the atom in its ground state since it has just emitted this photon. The probability of seeing a second photon at $\tau = 0$ is zero since the atom cannot re-radiate from the ground state. One must allow some time to elapse so that these may be a finite probability for the atom to be

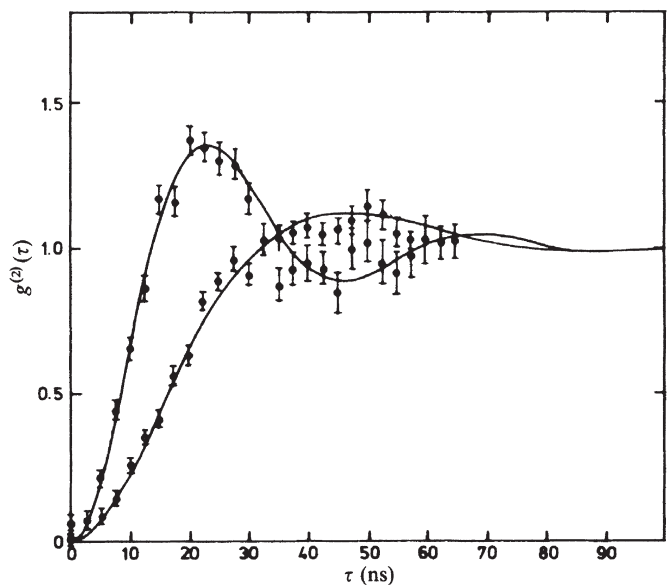


Fig. 4 Photon correlation measurements of fluorescent light from sodium. Experimental results obtained by Dagenais and Mandel compared with theory (solid line). ●, $\Omega/\gamma = 2.2$; ○, $\Omega/\gamma = 1.1$.

in the excited state and hence a finite probability for the emission of a second photon. In fact $g^{(2)}(\tau)$ is proportional to the probability that the atom will be in its excited state at time τ given that it was initially in the ground state.

Thus a single atom undergoing resonance fluorescence became a candidate for observing photon antibunching. However, the results (equations (7), (8)) hold only for resonance fluorescence from a single atom. If photons from many atoms contribute to the signal detected then one gets interference effects and the antibunching is diminished and for a large number of atoms lost entirely. In fact for a large number of independently contributing atoms one finds $g^{(2)}(\tau) = 2$ in agreement with the central limit theorem³.

It is clear therefore that an experiment of high sensitivity was necessary if measurements were to be made on the second order correlation function of light from a single atom. Such an experiment was performed by Kimble, Dagenais and Mandel⁴. A sketch of their experimental setup is shown in Fig. 3. In a similar configuration to the experiments which measured the spectrum they used an atomic beam of sodium atoms optically pumped to prepare a pure two-level system. The atomic beam was irradiated at right angles with a highly stabilised dye laser tuned on resonance with the $3P_{3/2}$, $F = 3$, $M_F = 2$ to $3^2S_{1/2}$, $F = 2$, $M_F = 2$ transition in sodium. The intensity of the atomic

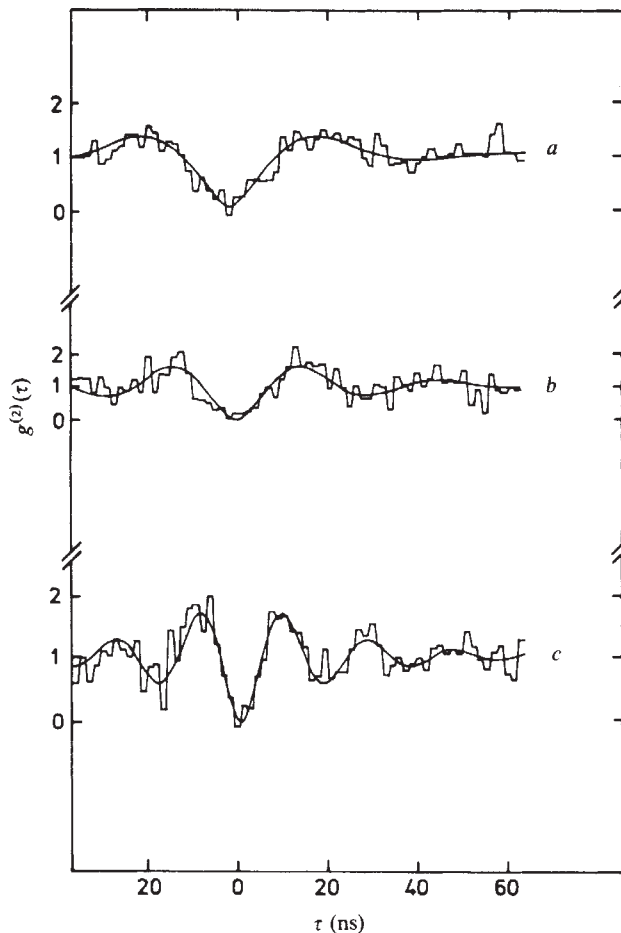


Fig. 5 Photon correlation measurements of fluorescent light from sodium. Experimental points adapted from the results of Leuchs *et al.* (personal communication) compared with theory. *a*, $\Omega/\gamma = 2.2$; *b*, $\Omega/\gamma = 4.3$; *c*, $\Omega/\gamma = 5.7$.

beam was reduced so that on the average no more than one atom is present in the observation region at a time. The fluorescent light from a small observation volume is observed in a direction orthogonal to both the atomic and laser beams.

The fluorescent light in this direction is divided into two equal parts by a beam splitter and the arrival of photons in each beam is detected by two photomultipliers. The pulses from the two detectors are fed to the start and stop inputs of a time to digital converter (TDC) where the time intervals τ between start and stop pulses are digitised in units of 0.5 ns and stored. The number of events $n(\tau)$ stored at address τ is therefore a measure of the joint photoelectric detection probability density which equals $\eta^2 g^{(2)}(\tau)$ where η is the detector efficiency. The initial results obtained by Kimble *et al.*⁴ showed the initial positive slope of $g^{(2)}(\tau)$ characteristic of photon antibunching but starting with $g^{(2)}(0) = 1$ rather than zero.

The reason for this disagreement with the theory was pointed out by Jakeman *et al.*⁴⁰ who attributed it to number fluctuations in the atomic beam. Though the atomic beam density is such that on the average only one atom is in the observation volume during a correlation time, there are poissonian fluctuations about this mean value so that at times there could be two or more atoms in the observation volume. Thus in the atomic beam experiment one observes the photon antibunching superimposed on the poissonian number fluctuations of the atoms. A calculation including the atomic number fluctuations was carried out by Carmichael *et al.*⁴¹ Similar calculations were made by Kimble *et al.*⁵ who obtained very good agreement with experimental results. Some recent experimental results of Dagenais and Mandel⁶ are shown in Fig. 4, where when the effects of

atomic number fluctuations are taken into account excellent agreement is obtained with the theory. A photon correlation experiment on resonance fluorescence from a beam of sodium atoms has been recently performed by Leuchs *et al.* (personal communication). The results of this experiment are shown in Fig. 5, where with compensation for the atomic number fluctuations excellent agreement is found with the theory. These experiments show clear evidence for the existence of photon antibunching and thus verify the predictions of the quantum theory of light.

Discussion

It could still be argued that one has not observed light with the antibunching property but only inferred from the experimental observation that the fluorescent light from a single atom must possess the antibunching property. This is due to the finite correlation at $\tau = 0$ arising from the number fluctuations of the atoms. It would be useful, therefore, to consider ways of improving experimental methods. Recently advances have been made towards isolating a single atom in an ion trap⁴². This suggests the possibility of performing resonance fluorescence on a single stationary atom which should radiate light with the antibunching character. In addition there are the suggested systems in nonlinear optics for example, sub-second harmonic generation²³⁻²⁶ and two-photon absorption²⁷⁻³³ where photon antibunching has been predicted. These experiments involve the nonlinear coupling of electromagnetic waves through a medium which may be in the solid state thereby eliminating problems due to atomic number fluctuations.

Although minor refinements may be possible the experiments performed by Kimble *et al.*⁴⁻⁶ and Leuchs *et al.* (personal communication) represent a fundamental contribution to our understanding of the nature of light. For the first time photon correlation experiments have detected an effect which is a direct manifestation of the quantum nature of light.

- Hanbury-Brown, R. & Twiss, R. Q. *Nature* **177**, 27 (1956); *Proc. R. Soc.* **242A**, 300 (1957); **243A**, 291 (1957).
- Glauber, R. J. *Phys. Rev.* **130**, 2529 (1963); **131**, 2766 (1963); in *Quantum Optics and Electronics*, (eds de Witt, C., Blandin, A. & Cohen-Tannoudji, C.) (Gordon and Breach, Edinburgh 1964).
- Carmichael, H. J. & Walls, D. F. *2nd Natn. Quantum Electronics Conf.* (1975); *J. Phys.* **9B**, L43 1199 (1976).
- Kimble, H. J., Dagenais, M. & Mandel, L. *Phys. Rev. Lett.* **39**, 691 (1977).
- Kimble, H. J., Dagenais, M. & Mandel, L. *Phys. Rev.* **18A**, 201 (1978).
- Dagenais, M. & Mandel, L. *Phys. Rev.* **18A**, 2217 (1978).
- Planck, M. *Verh. dt. Phys. Ges.* **2**, 202 (1900).
- Einstein, A. *Phys. Z.* **17**, 132 (1905); **18**, 121 (1917).
- Lamb, W. E., *Phys. Rev.* **128**, 885 (1947).
- Scitizky, I. R., Dente, G. C., Sachs, M. & Jaynes, E. T. in *Coherence and Quantum Optics*, Vol. 4 (eds Mandel, L. & Wolf, E.) (Plenum, New York, 1978).
- Taylor, G. I. *Proc. Cam. Phil. Soc. Math. Phys. Sci.* **15**, 114 (1909).
- Dirac, P. A. M. *Quantum Mechanics* (Oxford University Press, 1958).
- Walls, D. F. *Am. J. Phys.* **45**, 952 (1977).
- Rebka, G. A. & Pound, R. V. *Nature* **180**, 1035 (1957).
- Twiss, R. Q. & Little, G. A. *Aust. J. Phys.* **12**, 77 (1959).
- Morgan, B. L. & Mandel, L. *Phys. Rev. Lett.* **33**, 1397 (1964).
- Arecchi, F. T., Gatti, E. & Sona, A. *Phys. Lett.* **20**, 27 (1966).
- Cummins, H. Z. & Pike, E. R. (eds) *Photon Correlation and Light Beating Spectroscopy* (Plenum, New York, 1974).
- Haken, H. *Handbuch der Physik* Vol. XXV/2c (Springer, Berlin, 1970).
- Louisell, W. H. *Quantum Statistical Properties of Radiation* (Wiley, New York, 1973).
- Sargent, M., Scully, M. O. & Lamb, W. E. *Laser Physics* (Addison Wesley, New York, 1974).
- Sudarshan, E. C. G. *Phys. Rev. Lett.* **10**, 277 (1963).
- Stoler, D. *Phys. Rev. Lett.* **33**, 1397 (1974).
- Kielich, S., Kozierowski, M. & Tanas, R. in *Coherence and Quantum Optics IV* (eds Mandel, L. & Wolf, E.) (Plenum, New York, 1978).
- Mostowski, J. & Rzaewski, K. *Phys. Lett.* **66A**, 275 (1978).
- Drummond, P. D., McNeil, K. J. & Walls, D. F. *Opt. Commun.* **28**, 255 (1978).
- Chandra, N. & Prakash, H. *Phys. Rev. A1*, 1696 (1970).
- Tornau, N. & Bach, A. *Opt. Commun.* **11**, 46 (1974).
- Simaan, H. D. & Loudon, R. *J. Phys.* **8A**, 539 (1975).
- Bandilla, A. & Ritze, H. *Opt. Commun.* **19**, 169 (1976).
- Brunner, W., Mohr, U. & Paul, H. *Opt. Commun.* **17**, 145 (1976).
- Loudon, R. *Phys. Bull.* **27**, 21 (1976).
- Chaturvedi, S., Drummond, P. & Walls, D. F. *J. Phys.* **10A**, L187 (1977).
- Mollow, B. R. *Phys. Rev.* **188**, 1969 (1969).
- Wu, F. Y., Grove, R. E. & Ezekiel, S. *Phys. Rev. Lett.* **35**, 1426 (1975).
- Hartig, W., Rasmussen, W., Schieder, R. & Walther, H. *Z. Phys.* **A278**, 205 (1976).
- Schuda, F., Stroud, C. R. & Hercher, M. *J. Phys.* **B7**, L198 (1974).
- Cohen-Tannoudji, C. in *Frontiers in Laser Spectroscopy* (eds Batian, R., Haroche, S. & Liberman, S.) (North-Holland, Amsterdam, 1977).
- Kimble, H. J. & Mandel, L. *Phys. Rev.* **A13**, 2123 (1976).
- Jakeman, E., Pike, E. R., Pusey, P. N. & Vaughan, J. M. *J. Phys.* **10A**, L257 (1977).
- Carmichael, H. J., Drummond, P., Maystre, P. & Walls, D. F. *J. Phys.* **11A**, L121 (1978).
- Neuhauser, W., Hohenstatt, M., Toschek, P. & Dehmelt, H. *Phys. Rev. Lett.* **41**, 233 (1978).

Population biology of infectious diseases: Part II

Robert M. May

Biology Department, Princeton University, Princeton, New Jersey 08544

Roy M. Anderson

Zoology Department and Centre for Environmental Technology, Imperial College, London University, London SW7, UK

In the first part of this two-part article (Nature 280, 361–367), mathematical models of directly transmitted microparasitic infections were developed, taking explicit account of the dynamics of the host population. The discussion is now extended to both microparasites (viruses, bacteria and protozoa) and macroparasites (helminths and arthropods), transmitted either directly or indirectly via one or more intermediate hosts. Consideration is given to the relation between the ecology and evolution of the transmission processes and the overall dynamics, and to the mechanisms that can produce cyclic patterns, or multiple stable states, in the levels of infection in the host population.

IN the first part of this article¹ we considered the dynamics of microparasitic infections with direct transmission between hosts. We now extend the discussion to other kinds of parasites and transmission processes, and examine the general relations between population behaviour and parasite life cycle structure. The conclusions are broadly similar to those in the first part¹, but there are interesting similarities and differences both in the mathematical structure and in the biological conclusions.

We then give a brief discussion of general evolutionary trends, and end with a survey of the main mechanisms that can produce cyclic patterns, or multiple stable states, in the levels of infection in the host population.

Life cycle structure and disease dynamics

Macroparasites with direct life cycles tend to produce persistent infections, with the host harbouring populations of parasites for long periods, due to continual reinfections. Among many examples are the hookworm species of man, *Ancylostoma duodenale* and *Necator americanus* (see Table 1); in endemic areas the prevalence of these infections may approach² 100%. For such systems, the pathogenicity to the host, the rate of production of transmission stages of the parasite and any resistance of the host to further infection all typically depend on the number of parasites present in a given host. A crude division of the host population into susceptible, infected and immune classes is therefore not helpful, and a detailed description of the dynamics needs to deal with the full probability distribution of parasites within the host population^{3–6} (that is, with the number of hosts harbouring i parasites $N(i)$, where $i = 0, 1, 2, \dots$). Figure 1, which is to be compared with Fig. 3 of the first part of this article¹, depicts the essential structure of such models.

It is often useful to simplify these models by making a phenomenological assumption about the statistical distribution of parasites among hosts^{3,7–10} (or even, occasionally, by making assumptions that permit this distribution to be deduced theoretically^{11,12}). A usual phenomenological assumption is that the parasite distribution is a negative binomial^{3,7–10,13,14}, with the parameter k providing an inverse measure of the degree of parasite 'clumping' or overdispersion within the host population; the limit $k \rightarrow \infty$ corresponds to the parasites being distributed in an independently random or Poisson form, while very small k corresponds to very high clumping. It is then possible to use such statistical moments of the $N(i)$ distribution as the total host population ($N = \sum_i N(i)$), the number of uninfected hosts ($X = N(0)$), the total parasite population ($P = \sum_i iN(i)$), and the mean parasite burden per host ($m = P/N$). In

this way, models of the kind depicted in Fig. 1 can be brought into correspondence with the coarser models of the kind discussed in Part I (see Fig. 3 of Part I)¹.

The most detailed study of this type^{3,4} draws on a synoptic collection of data for direct life cycle parasites (mainly helminths), and describes the dynamics in terms of three differential equations, for the number of hosts N , parasites P , and free-living infective stages w :

$$dN/dt = (a - b)N - \alpha P \quad (1)$$

$$dP/dt = \beta wN - (\mu + b + \alpha)P - \alpha(k+1)P^2/(kN) \quad (2)$$

$$dw/dt = \lambda P - cw - \beta wN \quad (3)$$

Here the birth and death rates a and b are as defined in the first part¹ of this article, as is the transmission parameter β (hosts acquire individual adult parasites at a rate proportional to the number of contacts between hosts and parasite infective stages, βwN). The parasite-induced host death rate (or, equivalently, depression of the birth rate) is taken to be linearly proportional to the parasite burden in a given host, at a rate α per parasite. The parasites are distributed as a negative binomial with parameter k ; μ is the natural mortality rate of adult parasites; λ is the rate of production of infective stages by an adult parasite; and c is the death rate of these infective stages. The biological underpinning of these equations, and their dynamical behaviour, have been expounded in detail elsewhere^{3,4}.

A rough understanding of the relation between this system of equations for typical macroparasites with direct transmission, and the earlier set of equations (8)–(10) of Part I for directly transmitted microparasites, can be obtained as follows. First, note that the lifespan of the free-living infective stages is usually much shorter than that of the host and the adult parasite (compare Table 1). Thus the set of differential equations can be decoupled, by assuming the 'short lived' infective stages are adjusted essentially instantaneously to their equilibrium level ($dw/dt = 0$) for any given value of N and P . This gives

$$dN/dt = rN - \alpha P \quad (4)$$

$$dP/dt = \frac{\lambda NP}{H_0 + N} - (\mu + b + \alpha)P - \frac{\alpha(k+1)P^2}{kN} \quad (5)$$

(where $r = a - b$ and $H_0 = c/\beta$). Second, a phase-plane analysis now lays bare the properties of this pair of equations.

Three patterns of dynamical behaviour are possible^{3,4}. (1) If

$$\lambda - (\mu + b + \alpha) > r(k+1)/k \quad (6)$$

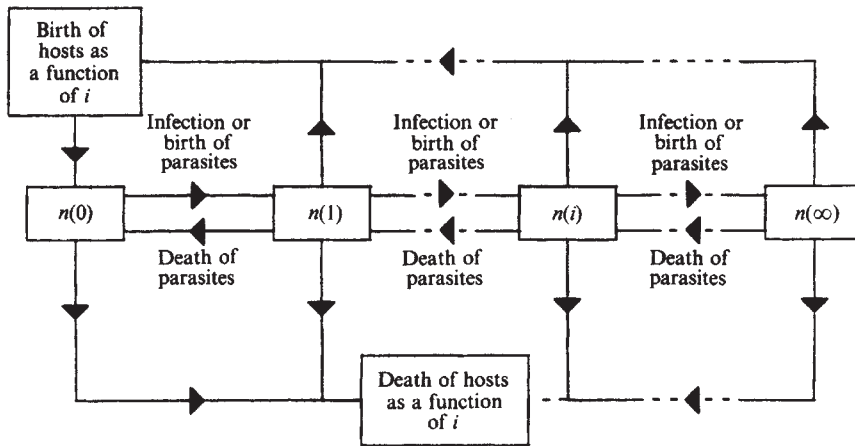


Fig. 1 Diagrammatic flow chart for a directly transmitted infection, based on a model with compartments for the number of hosts, $N(i)$, harbouring i parasites ($i = 0, 1, 2, \dots$). The model has a structure similar to, but obviously more complex than, that of Fig. 3 Part I¹.

the parasite regulates the host population to a stable equilibrium value. The average parasite burden per host settles to

$$m = r/\alpha \quad (7)$$

(2) If Equation (6) is not satisfied, but

$$\lambda - (\mu + b + \alpha) > 0 \quad (8)$$

the host population continues to grow exponentially, but at a rate

$$\rho = r - [\lambda - (\mu + b + \alpha)] [k/(k+1)] \quad (9)$$

This is less than the disease-free rate, r . In this case, the mean parasite burden in the exponentially growing host population settles to the value

$$m \rightarrow (r - \rho)/\alpha \quad (10)$$

In either event, if the host population is initially below the value

$$N_T = \frac{H_0(\mu + b + \alpha)}{\lambda - (\mu + b + \alpha)} \quad (11)$$

the parasite cannot become established ($dP/dt < 0$). However, as long as equation (8) is satisfied, the host population will grow exponentially (at the rate r) until this threshold value N_T is exceeded, whereupon the infection will become established, either regulating the host population or at least slowing its growth rate. Furthermore, in view of the large values for the reproductive output λ of most helminth parasites, N_T will typically be relatively small. This expectation of commonly finding direct life cycle helminth infections persisting in low density host populations is borne out by the evidence^{15,16}.

(3) Finally, if λ is so small that equation (8) is not satisfied, the infection can never become established (N_T is negative!).

The similarities between cases (1) and (2) here, and the results displayed in Fig. 4 of Part I, are striking. In particular, for measures of the prevalence of infection, notice the exact formal equivalence between equation (15) of Part I and equation (7), and between equation (17) of Part I and equation (10). A dissimilarity is that whereas the ability of a microparasitic

infection to regulate its host population essentially depends on its pathogenicity α exceeding the host population growth rate r (weighted by rates of recovery, loss of immunity and so on: see Table 1 of Part I), for a macroparasite it is its net reproductive ability, $\lambda - (\mu + b + \alpha)$, that plays a central role (λ is the 'birth rate', while μ , b and α are the natural parasite, natural host and parasite-induced host death rates). The macroparasitic infection can never persist if this effective net reproductive rate is not positive (equation (8)). The parasite will regulate the host population, or merely slow its growth, depending on whether this effective net reproductive rate $\lambda - (\mu + b + \alpha)$ is, or is not, greater than the host reproductive rate r , weighted by a factor $(k+1)/k$ to allow for the clumped distribution of parasites. Thus equation (6) is for these directly transmitted macroparasites the analogue of the microparasite equation (13) in Part I¹.

Indirect life cycles constitute another qualitatively different kind of complication, arising when the life cycle of the parasite involves one or more intermediate hosts. This happens for both microparasites (for example, the arthropod-borne viruses or arboviruses such as yellow fever or Rocky Mountain spotted fever; the protozoan malaria species) and macroparasites (for example, schistosomes, the filarial worms causing onchocerciasis, and other roundworms and flatworms that involve dipteran, molluscan and other intermediate hosts). Malaria and schistosomiasis in human populations are the two parasites whose transmission cycles have been most fully studied and each enjoys its own independent and growing literature, both empirical and theoretical (see Table 2). Their basic dynamical character is, however, in many respects common to all parasites with indirect life cycles.

If we adopt the approach of equations (8)–(10) and Fig. 3 discussed in part I, namely dividing the host population into susceptible, infected and immune categories, we will in the simplest case have a system of six differential equations: three for the primary host (alternatively referred to as the definitive host, or final host) classes X , Y , Z ; three for the intermediate host populations X' , Y' , Z' . All existing models, however, assume the total populations of both primary host ($N = X + Y + Z$) and the intermediate host ($N' = X' + Y' + Z'$) are constant, unaffected by the dynamics of the disease. This reduces the system to four equations. If, furthermore, immunity is either ignored or handled by specific assumptions about 'superinfection', the Z and Z' classes are effectively removed to give two coupled differential equations for the number of primary hosts Y , and of intermediate vectors Y' , that are infected. This, in essence, is the source of the classic Ross-Macdonald^{17,18} malaria equations, the Nasell-Hirsch¹⁹ schistosomiasis model, and the Dietz²⁰ arbovirus equations.

These equations have been subjected to various kinds of more refined treatment, including age structure^{20–24}, immunity and 'superinfection'^{17,21,23–26} and the use of several immunological categories of hosts²⁶ (intermediate between Fig. 3 of Part I and Fig. 1 of Part II). However, essentially all the existing work on

Table 1 Expected lifespans of the host and parasitic stages involved in the life cycle of *Schistosoma mansoni* and *Ancylostoma duodenale*

	Population	Lifespan (yr)
<i>S. mansoni</i> (refs 24, 29, 82)	Man (primary host)	50.00
	Adult parasite	5.00
	Infected snails (intermediate host)	0.10
	Cercariae	0.003
	Miracidia	0.0009
<i>A. duodenale</i> (ref. 2)	Man	50.0
	Adult parasite	1.0
	Free-living infective stage	0.1

indirectly transmitted parasites retains the assumption that the populations of host and intermediate vector are constant, not dynamically involved with the infection. Analysis of such models reveals threshold relations^{17,20,21,27,28} between N and N' , analogous to but more complicated than the N_T of the direct life cycle models. If N and N' lie below the threshold combination, the disease cannot be maintained.

For many human, and other animal, infections by parasites with indirect life cycles, what is needed is a theory in which the populations of primary and intermediate hosts are affected, and possibly even determined, by the presence of the infection. While it may often be reasonable to treat a human primary host population as roughly constant, we believe that cases where intermediate host populations are unaffected by the prevalence levels of the infection will be the exception rather than the rule²⁹. There is no formal problem in extending our dynamic models of either the 'microparasite' kind of equations (8)–(10), (Part I) or the 'macroparasite' kind of equations (1)–(3) (Part II), to encompass the added complication of one or more intermediate vector populations. Space forbids a full exposition of the emergent properties, but the main trends are indicated in the following section.

Time scales and transmission terms

A full model for an indirectly transmitted parasite might include not only dynamical descriptions of the prevalence of infection in primary and intermediate host populations, but also additional differential equations (analogous to equation (3)) for the free-living transmission stages that carry the parasite from primary to intermediate host, and back again. For example, for schistosomiasis we could add a differential equation describing the miracidial stage (man to snail), and another for the cercarial stage (snail to man), to the usual equations for infection levels in the human and snail populations³⁰. The reason this is not commonly done can be seen from Table 1; the dynamics of the free-living stages takes place on a time scale so much shorter than the other time scales in the system that miracidial and cercarial populations can be assumed to have the equilibrium values appropriate to the prevailing conditions among human and snail populations. In just this way, we collapsed the three-equation system (1)–(3) to the two equations (4), (5).

This technique of using biological insights about the time scales of various infection processes can be used to make further rough but useful approximations. For example, the time scales for processes (such as mortality rates) within the intermediate host population are typically significantly shorter than those in the primary host. Again, Table 1 testifies to this. Accordingly, we can assume that the numbers of susceptible, infected and immune intermediate hosts are adjusted to have the equilibrium values ($dX'/dt = 0$, and so on) appropriate to the current levels

of infection in the primary hosts. In this way, parasitic infections with indirect life cycles can be approximately brought to a form similar to that of equations (8)–(10) in Part I for direct life cycles¹.

As a concrete example, consider a grossly oversimplified model for malaria, in which 'superinfection'^{17,25,26}, and mosquito latency^{17,21} (and immunity^{20,26}), are ignored. Assume also the total mosquito population is constant; $N' = X' + Y' = \text{constant}$. The populations of infected humans Y , and mosquitoes Y' , then obey

$$dY/dt = \beta' Y' X - (b + \alpha + v) Y \quad (12)$$

$$dY'/dt = \beta Y(N' - Y') - (b' + \alpha' + v') Y' \quad (13)$$

Here β , b , α and v (plain for humans, primed for mosquitoes) have their previous meanings; conventionally, most infected humans are assumed to recover ($v \gg \alpha, b$), and most infected mosquitoes to die at a rate largely unaffected by the infection ($b' \gg \alpha', v'$). The assumption that mosquito processes happen on a relatively fast time scale enables Y' in equation (12) to be determined by setting $dY'/dt = 0$ in equation (13), leading to

$$dY/dt = Y[(\beta\beta'N'X)/(b' + \alpha' + v' + \beta Y) - (b + \alpha + v)] \quad (14)$$

This is exactly of the form for a directly transmitted infection (equation 9 of Part I), except that the simple transmission coefficient β has been replaced by the more complicated factor $\beta\beta'N'/(b' + \alpha' + v' + \beta Y)$. Similarly, the Nasell-Hirsch two equation model¹⁹ for prevalence of schistosomiasis among humans and snails can be collapsed to Macdonald's^{27,28} single equation for prevalence in the human population.

Conversely, for humans the total population is often growing on a time scale that is long compared to the relevant time scales of even persistent infections. This is why the total population can be treated as a constant in most epidemiological models. The approximation, whereby the dynamics of the prevalence (Y/N) and of the total population (N) are decoupled, can often be useful in discussing the transmission cycle of the infection, even though the long-term growth or regulation of the host population is affected by the presence of the infection.

Table 2 uses these ideas to attempt to give a schematic account of the relations among some of the many models, of differing degrees of complexity, that are to be found in the literature.

Saturation of transmission terms. The transmission terms obtained in equation (14) by 'collapsing out' the mosquito dynamics of equation (13), and in equation (5) by collapsing equation (3) for the free-living infective stages of the parasite, manifest a feature that is common to all such approximate representations of complex transmission processes^{3,4,14,31,32}. Essentially, the simple term βXY for direct transmission

Table 2 Schematic representation of relationships between various kinds of models for parasitic infections, based on relative time scales of population processes

	Host population(s) constant	Host population(s) a dynamic variable
In considering the dynamics of infection, only one species is involved (for example, the host species).	Direct life cycles. Classical epidemiological models (refs 21, 56, 65, 67, 83–84, 87). Models for the dynamics of a parasite population within a host population of fixed size (refs 5, 88, 89). Indirect life cycles. Models for schistosomiasis (refs 28, 34, 90) and for malaria (refs 17, 18, 21), considering only the dynamics within the human host.	Direct life cycles. Models similar to those for classical epidemiology, but with total host population a dynamic variable, determined by birth and death processes (refs 35, 91 and this review). Direct and indirect life cycles. Dynamics of models in the compartment below but with all populations but the primary host 'collapsed out' (this review).
In considering the dynamics of infections, two or more species are involved (e.g., primary and intermediate host, or host and parasite population).	Direct life cycles. Models similar to the classical epidemiological equations, but including the dynamics of free-living infective stages (ref. 35). Indirect life cycles. Models for schistosomiasis (refs 19, 34, 92), malaria (refs 17, 21, 93) and arbovirus infections (ref. 20) in which both human hosts and intermediate vectors are considered. Models of schistosomiasis where humans, snails, miracidiae and cercariae are all considered (ref. 30).	Direct life cycles. Models similar to classical epidemiology, but with host population and free living infective stages both included as dynamic variables (ref. 35). Models for dynamics of host parasite systems (refs 3, 4, 7, 8, 14, 35), sometimes with dynamic aspects of free-living infective stages also included (refs 3, 4). Indirect life cycles. Any of the models in the compartment to the left, but with the total host populations treated as dynamic variables (this review).

between susceptible and infected people or βNw for direct transmission between hosts and free-living infective stages are replaced by expressions of the general form $AXY/(1+\nu Y)$ or $ANP/(1+\nu N)$, respectively. In the limit when, for example, νY is small, the expression has the familiar form, proportional to X and Y . But it can be that νY becomes significant compared to unity, whereupon the transmission term saturates to a value (AX/ν) proportional only to X . Such saturation effects can be important in diminishing the ability of the parasite to regulate its host population^{3,14}.

Ecology of the transmission process. Further complications can arise from the ecological nature of the individual links in the transmission process.

For infections that are communicated directly, the assumption that the net rate is proportional to the number of susceptibles and to the number of infectives is clearly reasonable for many diseases, and strikingly successful in explaining the mouse pox and mouse pasteurellosis data¹. But for sexually transmitted diseases, for example, this is only plausible in a population that is astonishingly promiscuous and sexually active. In a society whose members typically have only a small number, η , of sexual partners (independent of the absolute population size), the rate at which an infected person propagates the infection is proportional not to the total number of susceptibles, but to η times the probability that a given person is a susceptible; that is, βXY is to be replaced²¹ by $\beta\eta XY/N$. Under these very simple assumptions, the condition for maintenance of such diseases is $\beta\eta > (b + \alpha + \nu)$, independent of the population size. In reality, a more careful treatment of the distribution of degrees of sexual activity within the population is needed³³, but the fact remains that infections of this sort are relatively easy to maintain in low density populations.

More broadly, biological insights into the relative time scales associated with the various phases of indirect life cycles enable us to discuss the prevalence of infection in the primary host population by retaining equations (8) and (10) for X and Z , in Part I of this article, but replacing equation (9) with the more general expression

$$dY/dt = Y[h - (a + b + \nu)] \quad (15)$$

The transmission term is here denoted by h (Ross' 'happenings'¹⁸), and the threshold condition for the disease to increase upon introduction at low levels is clearly that $h > (a + b + \nu)$ in the limit $Y \rightarrow 0$. For the simple circumstances of the indirect life cycle that led to equation (14) above, this requirement comes down to the threshold criterion^{20,21}

$$NN' > \frac{(\alpha + b + \nu)(\alpha' + b' + \nu')}{\beta\beta'} \quad (16)$$

Note that a large population N' of intermediate vectors can enable the disease to persist, even when the primary host population N is small.

However, for malaria and many other infections borne by biting arthropods, the intermediate vector tends to make a fixed number of bites per week, independent of the number of primary hosts available to feed on. Thus the transmission rate from infected arthropods to people (and from infected people back to susceptible arthropods) is proportional to the biting rate ω times the probability that a given human is susceptible (or infected), and not simply proportional to the number of susceptible (or infected) people. That is, in equations (12) and (13), β and β' are to be replaced by ω/N . The threshold condition (16) is accordingly modified to^{17,18,20}

$$\frac{N'}{N} > \frac{(\alpha + b + \nu)(\alpha' + b' + \nu')}{\omega^2} \quad (17)$$

Note that latency effects have been neglected here, although they can be important in infections with indirect life cycles, and they certainly modify threshold conditions significantly for malaria¹⁷ and schistosomiasis²⁹. Infections with intermediate

vectors of this character are relatively easy to maintain at low population densities of the primary host, provided only that the ratio of intermediate to primary hosts is sufficiently high. Indeed, equation (17) suggests the infection is actually easier to maintain at low host population levels; the mosquito or other intermediate host population N' is, however, typically dependent on the primary host for blood meals or the like, so that things are not as simple as they might seem. (A more general discussion, from which the threshold relations (16) and (17) emerge as limiting cases, has been given by Dietz²⁰).

Yet another form of complication enters with parasites that have sexual stages, yet can have low densities, in a host. Schistosomiasis is one such example^{10,19,27,28,34}. At high levels of prevalence of the infection in the human population, people tend to have worm burdens such that most adult female schistosomes are mated, and the circumstances leading to equation (16) are well approximated. But at low levels of prevalence, it can be that the average female is not mated, which tends to require that the transmission link from snail to man be counted twice in considering the overall cycle, thus giving complicated threshold conditions (very roughly of the form $N[N']^2 > \text{constant}$ ^{19,34}).

Finally, note that (apart from the laboratory experiments on mice¹) in all our models the host population either is regulated to some stable value by the disease, or else it grows exponentially. In practice, other constraints, set by resources, predators or the like, will eventually limit population growth. Such biological realities can be included in all our models, by introducing a logistic constraint (at a 'carrying capacity' K) in the growth of the disease-free population³⁵. The resulting situation, for both direct and indirect parasite life cycles, is similar to that illustrated in Figs 1e, f and 2b in Part I, with the host population depressed below its disease-free level K , provided the parasite-induced host mortality α is not too large¹. Too small an α leads to relatively little depression of the host population; too large an α renders the disease unable to persist, and the host population remains at K ; maximum parasite-induced depression of the host population is attained for intermediate levels of pathogenicity^{1,36}. This broad statement glosses over many intricacies that can arise (R.M.M. and R.M.A., in preparation), particularly with indirect life cycles when the intermediate vector has a constant biting rate (producing threshold conditions such as equation (17) in simpler models), but the gist is true.

Population parameters and evolutionary trends

Any discussion of the relations among the population parameters that characterize an infectious disease must ultimately take account of the evolutionary pressures on both hosts and parasites. Population dynamics is always confounded by population genetics.

For example, even if we assume no genetic change in the parasite, its action on the host will select for individuals with reduced susceptibility to the disease. For this reason alone, the pathogenicity of the parasite will tend to decrease through evolutionary time. Conspicuous examples are provided by the presence of the sickling gene (and other blood-group phenomena) in regions where malaria is endemic³⁷, and by the history of myxomatosis in rabbit populations in Australia³⁸. An interesting theoretical discussion has been given by Gillespie³⁹.

Selective forces also act strongly on the parasites^{40,41}. As we have seen, the persistence of a disease is facilitated by low pathogenicity and by long duration of infection¹. Countervailing forces can, however, act to increase the virulence of an infectious disease; increased pathogenicity may often be associated with enhanced rates of production (within the host) of the parasite's transmission stages^{38,42,43}.

The regulatory potential of an infectious disease will, therefore, typically change as time goes by. A parasite may stably regulate its host population during their early association. But,

as selective pressures reduce the average susceptibility of the hosts, such regulatory effects will tend to wane. Eventually, the host population may escape being controlled by the parasitic infection.

Because the generation times of most hosts are several orders of magnitude longer than those of their parasites, it is tempting to conclude that selection acts more rapidly on the parasites. However, the way parasitic infections act within host populations makes it likely that the parasites force the pace of host evolution to keep in step with, or even ahead of, their own evolution.

Among the recondite variety of strategies that parasites have evolved for persistence and transmission, some general trends can be discovered. For example, many parasitic species traverse links in community food webs by virtue of predator-prey associations between primary and intermediate hosts. Such associations, which include biting arthropods feeding on vertebrates, have played an important part in the evolution of complex life cycles. The high transmission efficiency β of these links suggest the threshold host populations for maintenance of such parasites will be low (see equations (16) and (17)). Consequently, we expect indirect life cycles to predominate among parasitic infections of hosts that exist at low density.

In contrast, directly transmitted microparasites that require high host densities in order to persist should be more commonly associated with animals that exhibit herd or shoaling behaviour, or breed in large colonies. Empirical evidence in support of these ideas comes from the abundance of directly transmitted viral and bacterial infections within modern human societies^{42,44}, large herds of ungulates⁴⁵, breeding colonies of sea birds^{46,47}, and the social insects^{48,49}. Those diseases with direct life cycles that do persist within low density host populations should possess distinctive characteristics, such as long-lived infective stages^{50,51}, failure to induce lasting immunity^{33,52-53}, or ability to persist within the host for very long times⁵⁴.

Another trend to be noted is that highly pathogenic species usually exist, if at all, at low levels of prevalence (see equations (15) and (17) in Part I and equations (7) and (10) in Part II). An example is the digenetic parasite *Haematolaecheus colaradensis* whose primary host is frogs, but which has a transmission pathway involving first snails, and then dragonflies, as intermediate vectors on the way to the next frog. The prevalence among frogs is high, 60–70%, and the parasite is long-lived and has very low pathogenicity; in dragonflies the fluke induces moderate mortality, and has 30–40% prevalence; while in snails it is highly pathogenic but appears to have only about 5–10% prevalence⁵⁵. These broad patterns, which are often found in helminths with life cycles involving two or three host species, are summarized schematically in Table 3.

Cyclic patterns of disease prevalence

Annual or other cycles in the prevalence of infection are often observed, and can arise in at least three distinct ways.

First, for many short-lived viral and bacterial infections in human populations, there is a propensity for the steady, endemic level of prevalence of infection to be attained by damped oscillations. Particularly if this equilibrium prevalence is low, it is possible for stochastic fluctuations in the number of people infected at the minimum of the cycle to, in effect, keep the cycle

'pumped' and prevent it from damping to equilibrium. This interplay between demographic stochasticity and an inherent propensity to weakly damped oscillations is essentially the mechanism proposed in the classic work of Bartlett^{21,56-59} to account for cyclic patterns in the prevalence of measles and other viral infections in large cities. Gurney and Nisbet⁶⁰ have proposed a similar mechanism as an explanation for predator-prey cycles.

Second, time dependence in any of the population parameters may, in principle, produce cyclic variations in infection. In particular, seasonal variation in the transmission coefficient is important in setting temporal patterns for many parasitic infections, and may often be central for human viral infections^{44,61-65}. The mechanisms underlying the seasonality in the transmission rates are poorly understood, but for human viruses the main causes are probably climatic (temperature and humidity) effects influencing survival and dispersal of transmission stages, and seasonal changes in social behaviour^{66,67} (children returning to school after the long summer vacation). The seasonal cycles characteristic of the prevalence of measles, chicken pox, poliomyelitis and mumps in large cities could arise in this way^{21,44,65}.

Annual periodicity in transmission rates can, moreover, produce complicated nonseasonal cycles in the prevalence of infection. Yorke and co-workers^{44,63,65}, and Dietz⁶⁴, have cogently argued that such a mechanism is responsible for the regular biennial cycle, alternating between years of high and low incidence, for measles in New York City between 1948 and 1964; in the same city, mumps and chicken pox showed clear annual cycles. The explanation of Yorke *et al.* is to the contrary of the conventional explanation of these non-seasonal cycles in terms of demographic stochasticity, as described above. Their model is essentially the set of deterministic equations (1)–(3) in Part I, with an assumed constant number of new susceptibles appearing each year, life-long immunity, and with the system enriched by the inclusion of a brief incubation period during which infected hosts are not infective. The basic feature is that the transmission coefficient $\beta(t)$ varies seasonally with a 1 yr period. Within a narrow window of parameter values, the number of infected people can show biennial peaks (Fig. 2a) similar to those for measles in New York City. This window separates highly transmissible diseases which produce an epidemic with eventual fade-out, from the diseases with low transmission efficiency which give rise to endemic seasonal patterns of infection (Fig. 2b), as usually shown by mumps and chicken pox.

Third, various kinds of nonlinearities in the transmission terms may produce stable limit cycles, whose periods depend on the population parameters and will rarely be seasonal. People familiar with the ease whereby stable limit cycles arise in predator-prey models may be surprised to learn the structure of most host-parasite models is such that stable cycles do not easily occur. However, they can be produced without excessive contrivance. One simple example is to take the basic equations (8)–(10) of part I, and introduce the possibility of saturation in the transmission by replacing βXY by $\lambda XY/(H_0 + X)$. Such a modification can arise naturally^{4,35}, in the manner of the analogous expression in equations (5) and (14), if the term is thought of as deriving from the 'collapsed' dynamics of a free-living infective stage. This system can now exhibit stable limit cycles for a specific range of parameter values (corresponding to λ neither too small nor too large). One such stable cycle is illustrated in Fig. 2c. In general, however, little is yet agreed about the kinds of biological processes that can generate nonseasonal patterns of disease prevalence.

Multiple stable states of disease prevalence

A growing number of empirical and theoretical studies suggest that many natural assemblies of plants and animals can have a multiplicity of alternative stable states⁶⁸. Once two or more stable states are possible, the actual state the system settles into

Table 3 Some population characteristics of diseases caused by indirectly transmitted helminths

Host	Pathogenicity of parasite	Prevalence of infection within host population	Expected life span of host (inversely related to time scaled dynamics)
Final	Low	High	Long
Second intermediate	Medium	Medium	Medium
First intermediate	High	Low	Short

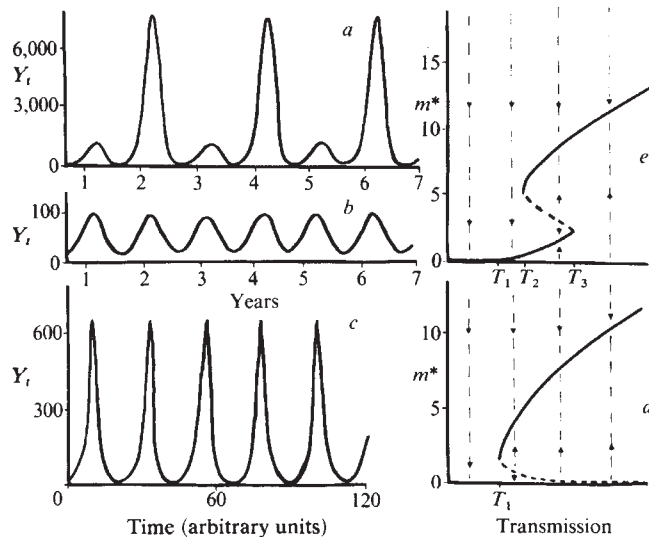


Fig. 2 *a*, Simulations of recurrent outbreaks of measles in New York City, showing biennial peaks superimposed on an underlying seasonal cycle (from London and Yorke⁴⁴). *b*, Simulations^{44,63} of recurrent outbreaks of mumps in New York City, with simple seasonal peaks, annually. *c*, Simple limit cycle behaviour generated by the model described in the text. *d*, The transmission threshold, alternative stable states, and 'breakpoint' phenomena that arise in simple models for the transmission dynamics of schistosomiasis^{10,19,27,28,34}; the features are as discussed in the text. *e*, Transmission threshold and alternative stable states arising in a model for directly transmitted helminth infections, where it is assumed that the pathogenicity of the disease is related to the nutritional state of the host³⁵. The graph shows the mean equilibrium burden of parasites per host, m^* , as a function of a parameter, T , representing transmission efficiency. The infection cannot persist below a threshold value T_1 ; between T_1 and T_2 there is a unique low level of disease endemicity; between T_2 and T_3 two stable levels of prevalence may occur, one high and the other low, separated by a breakpoint (the dashed line); above T_3 there is again a unique equilibrium level, corresponding to high average parasite burdens per host. The arrows indicate the stable state to which the system will converge from a given initial value.

depends on the initial conditions. The system will tend to recover its original configuration if subject to small disturbance, but sufficiently severe perturbations are liable to precipitate it into an alternative state in a different region of the dynamical landscape.

The nonlinearities in population models for parasitic infections can generate such multiple states by three principal mechanisms: worm pairing for sexual reproduction in the primary host; nonlinearities associated with the transmission from primary to intermediate host, or vice versa (mosquitos biting man for malaria, or predatory primary hosts consuming infected intermediate-host prey); parasite pathogenicity dependent on the nutritional state of the host.

The first and most fully studied of these categories arises for many helminth infections with indirect life cycles, such as schistosomes^{10,19,27,28,34}. It serves to exemplify the phenomenon. As portrayed in Fig. 2*e*, the equilibrium value of the mean parasite burden per human host (m) will be zero if the rate of transmission (T) from snail to man is below the threshold value T_1 . Above this threshold, two alternative stable states occur, one of endemic infection ($m > 0$), the other of parasite absence ($m = 0$). The basic reason is that at low levels of m the female worms are unlikely to be mated, so that the disease cannot be maintained, even though the transmission parameters are such

as to permit its endemicity if introduced at high values of m . The two stable states (valley bottoms in the dynamical landscape) are separated by a 'breakpoint' (watershed), indicated by the dashed line in Fig. 2*e*; disturbances severe enough to transgress the breakpoint will carry the system from one state to the other.

These threshold and breakpoint concepts are of obvious importance to epidemiologists concerned with disease eradication^{10,27,28}.

Of special importance are the effects that can arise from the now widely recognised fact that the impact of an infection is often related to the nutritional state of the host⁶⁹⁻⁷⁵. Broadly speaking, malnourished hosts have lowered immunological competence, and are less able to withstand the onslaught of infection⁷⁶⁻⁷⁸. The effective pathogenicity of a parasite therefore tends to increase as host density rises to a level where competition for available food resources is severe^{79,80}. Given certain reasonable assumptions³⁵ about the exact relation between pathogenicity (α) and host density (N), two stable states may occur for a given set of rate parameters. The outcome of such a model³⁵, for a directly transmitted helminth infection, is shown in Fig. 2*d*. Both states reflect stable endemic disease: one equilibrium is characterized by high host density and low worm burdens; the other by low host density (severely depressed by the disease) and high average burdens of parasites. As for the schistosome model of Fig. 2*e*, the two states are separated by a breakpoint or unstable equilibrium.

The discontinuous switch from low to high levels of infection, following a disturbance severe enough to cross the breakpoint, will show up as an apparent 'epidemic' outbreak of disease, typically producing many host deaths. Interestingly, many documented accounts of disease outbreaks are for host populations at high densities, where stress induced by overcrowding or malnutrition is present^{71,72}. It is very likely that such outbreaks are to be explained³⁵ by the alternative stable states produced by close links between pathogenicity and nutrition or stress, rather than by the commonly accepted hypothesis of enhanced transmission with high density populations⁸¹.

Parasitic infections with very complex life cycles may possess more than two stable states, particularly if predator-prey links are involved in the transmission from one host to the next, as is the case for many helminth parasites. There is a desperate paucity of data, from field or laboratory, bearing on these general points.

Conclusion

This two-part article has blended some new theoretical studies and new analysis of existing laboratory data with a review and synthesis of past and present models for the overall transmission dynamics of parasitic infections. We have defined 'parasite' broadly to include viruses, bacteria and protozoans along with the more conventional helminth and arthropod parasites, and we have concentrated attention upon the circumstances under which the infection may significantly alter the growth rate of its host population.

Some of the theoretical conclusions can be pleasingly supported by hard data, while others remain more speculative. On the whole, our main goal is to help elevate the study of host-parasite population dynamics to its proper place in ecological thinking; parasites (broadly defined) are probably at least as important as the more usually-studied predators and insect parasitoids in regulating natural populations.

We are grateful to many people, and particularly to Mary Anderson, David Bradley and James Yorke, for their help. This work was supported in part by the NSF, under grant DEB77-01565.

- Anderson, R. M. & May, R. M. *Nature* **280**, 361-367 (1979).
- Hoagland, K. E. & Schad, G. A. *Exp. Parasit.* **44**, 36-49 (1978).
- Anderson, R. M. & May, R. M. *J. Anim. Ecol.* **47**, 219-247 (1978).
- May, R. M. & Anderson, R. M. *J. Anim. Ecol.* **47**, 249-267 (1978).
- Anderson, R. M. in *Ecological Stability* (eds Usher, M. B. & Williamson, M. H.) (Chapman and Hall, London, 1974).
- Fine, P. E. M. *Ann. N.Y. Acad. Sci.* **266**, 173-194 (1975).

- Crofton, H. D. *Parasitology* **63**, 179-193 (1971).
- Crofton, H. D. *Parasitology* **63**, 343-364 (1971).
- May, R. M. *J. Anim. Ecol.* **47**, 833-844 (1978).
- Bradley, D. J. & May, R. M. *Trans. R. Soc. trop. Med. Hyg.* **72**, 262-273 (1978).
- Tallis, G. M. & Leyton, M. K. *Math. Biosci.* **4**, 39-48 (1969).
- Anderson, R. M. in *Mathematical Models in Medicine* (eds Berger, J. et al.) (Springer, New York, 1976).

13. Anderson, R. M. *Parasitology* **76**, 119–157 (1978).
14. May, R. M. *Parasitology* **75**, 259–276 (1977).
15. Baruzzi, R. G., Macopito, L. F., Serra, M. L. C., Souza, F. A. A. & Stabile, C. in *Health and Disease in Tribal Societies* (CIBA Fdn. Symp. **49**) (North-Holland, Amsterdam, 1977).
16. Campbell, A. D. *Proc. R. Soc. Edinb.* **B74**, 347–369 (1972).
17. Macdonald, G. *The Epidemiology and Control of Malaria* (Oxford University Press, 1957).
18. Ross, R. *Br. med. J.* **i**, 546–547 (1915).
19. Nasell, I. & Hirsch, W. M. *Commun. pure appl. Math.* **26**, 395–453 (1973).
20. Dietz, K. in *Epidemiology* (eds Ludwig, D. & Cooke, K. L.) 104–121 (Society for Industrial and Applied Mathematics, Philadelphia, 1975).
21. Bailey, N. T. J. *The Mathematical Theory of Infectious Diseases* 2nd edn (Macmillan, New York, 1975).
22. Hairston, N. G. *Bull. Wild Hlth Org.* **33**, 163–175 (1965).
23. Cohen, J. E. in *Theoretical Ecology: Principles and Applications* (ed. May, R. M.) (Blackwell, Oxford, 1976).
24. Cohen, J. E. *A. Rev. Ecol. Syst.* **8**, 207–233 (1977).
25. Fine, P. E. M. *Proc. R. Soc. Med.* **68**, 547–551 (1975).
26. Dietz, K., Molineaux, L. & Thomas, A. *Bull. Wild Hlth Org.* **50**, 347–357 (1974).
27. Macdonald, G. *Dynamics of Tropical Disease* (eds Bruce-Chwatt, L. J. & Glanville, V. J.) (Oxford University Press, 1973).
28. Macdonald, G. *Trans. R. Soc. trop. Med. Hyg.* **59**, 489–506 (1965).
29. Anderson, R. M. & May, R. M. *Parasitology* (in the press).
30. Lewis, T. *Math. Biosci.* **30**, 205–210 (1976).
31. Macdonald, G. *Public Health Reports Washington D.C.* **76**, 753–764 (1961).
32. Keymer, A. E. & Anderson, R. M. *Parasitology* (in the press).
33. Yorke, J. A., Hethcote, H. W. & Nold, A. J. *sex. Trans. Dis.* **5**, 51–56 (1978).
34. May, R. M. *Math. Biosci.* **35**, 301–343 (1977).
35. Anderson, R. M. in *Population Dynamics* (eds Anderson, R. M., Turner, B. D. & Taylor, L. R.) (Blackwell, Oxford, 1979).
36. Anderson, R. M. *Nature* **279**, 150–152 (1979).
37. Bradley, D. J. in *Origins of Pest, Parasite, Disease and Weed Problems* (eds Cherrit, J. M. & Sagar, G. R.) (Blackwell, Oxford, 1977).
38. Fenner, F. & Ratcliffe, F. N. *Myxomatosis* (Cambridge University Press, 1965).
39. Gillespie, J. H. *Ecology* **56**, 493–495 (1975).
40. Price, P. W. *Evolution* **31**, 405–420 (1977).
41. Price, P. W. *Evolutionary Biology of Parasites* (Princeton University Press, 1979).
42. Black, F. L. *et al.* *Am. J. Epidem.* **100**, 230–250 (1974).
43. Eshel, I. *Theor. Pop. Biol.* **11**, 410–424 (1977).
44. London, W. P. & Yorke, J. A. *Am. J. Epidem.* **98**, 453–468 (1973).
45. Davis, J. W., Karstad, L. H. & Trainer, D. O. (eds) *Infectious Diseases of Wild Mammals* (Iowa State University Press, Ames, 1970).
46. Kaschula, V. R. & Truter, D. E. *J. S. Afr. vet. med. Ass.* **22**, 191–192 (1951).
47. Warner, R. E. *Condor*, 101–120 (1968).
48. Woodrow, A. W. *J. Econ. Ent.* **35**, 892–895 (1942).
49. Bailey, L. *Sci. Progr. (Oxford)* **59**, 309–323 (1971).
50. Tinsley, T. W. *A. Rev. Ent.* **24**, 63–87 (1979).
51. Henry, J. E. & Oma, E. A. *J. Invert. Path.* **23**, 371–377 (1974).
52. Maramorosch, K. *Invertebrate Immunity* (Academic, New York, 1975).
53. Smith, K. M. *Virus-Insect Relationships* (Longman, London, 1976).
54. Brook, T. D. *Microbial Ecology* (Prentice Hall, New Jersey, 1966).
55. Matumota, M. *Bact. Rev.* **33**, 404–418 (1969).
56. Bartlett, M. S. *Stochastic Populations in Ecology and Epidemiology* (Methuen, London, 1960).
57. Dronen, N. O. *Am. Midl. Nat.* **99**, 330–349 (1978).
58. Bartlett, M. S. *J. R. Stat. Soc. A120*, 48–70 (1957).
59. Bartlett, M. S. *J. R. Stat. Soc. A123*, 37–44 (1960).
60. Nisbet, R. M. & Gurney, W. S. C. *Nature* **263**, 319–320 (1976).
61. Anderson, R. M. *Parasitology* **72**, 281–305 (1976).
62. McNamara, M. J., Pierce, W. C., Cranford, Y. E. & Miller, L. F. *Am. Rev. resp. Dis.* **86**, 485–492 (1962).
63. Yorke, J. A. & London, W. P. *Am. J. Epidem.* **98**, 469–482 (1973).
64. Dietz, K. in *Mathematical Models in Medicine* (ed. Levin, S. A.) (Springer, New York, 1976).
65. Yorke, J. A., Nathanson, N., Pianigiani, G. & Martin, J. *Am. J. Epidem.* **109**, 103–123 (1979).
66. Bliss, C. I. & Blevins, D. L. *Am. J. Hyg.* **70**, 328–334 (1959).
67. Soper, H. E. *J. R. Stat. Soc. A92*, 34–73 (1929).
68. May, R. M. *Nature* **269**, 471–478 (1977).
69. Brooke, M. M. *Am. J. Hyg.* **41**, 81–108 (1945).
70. Cole, T. J. & Parkin, J. M. *Trans. R. Soc. trop. Med. Hyg.* **71**, 196–198 (1977).
71. Steinhaus, E. A. *Proc. 10th Int. Cong. Ent.* **4**, 725–730 (1958).
72. Geiman, Q. M. *Vitam. Horm.* **16**, 1–33 (1958).
73. Gibson, T. E. *Proc. nutri. Soc.* **22**, 15–20 (1963).
74. Chandler, A. M. C. *J. Egypt. med. Ass.* **36**, 533–552 (1953).
75. Wiger, R. *Oikos* **29**, 598–606 (1977).
76. Mimms, C. A. *The Pathogenesis of Infectious Diseases* (Academic, London, 1977).
77. Roitt, I. M. *Essential Immunology* (Blackwell, Oxford, 1976).
78. Scrimshaw, N. S., Taylor, C. E. & Gordon, J. E. *Wild. Hlth Ord. Monogr. Ser.* **57**, (1968).
79. Gordon, M. H. *Proc. Aust. Anim. Prod.* **3**, 93–104 (1963).
80. Sheppe, W. A. & Adams, J. R. *J. Parasit.* **57**, 55–59 (1957).
81. Lack, D. *The Natural Regulation of Animal Numbers* (Oxford University Press, 1954).
82. Warren, K. S. *J. infect. Dis.* **127**, 595–609 (1973).
83. Dietz, K. *J. R. Stat. Soc. A130*, 505–528 (1967).
84. Waltman, P. *Deterministic Threshold Models in the Theory of Epidemics* (Springer, New York, 1974).
85. Hoppens, F. C. *Mathematical Theories of Populations: Demographics, Genetics and Epidemics* (SIAM, Philadelphia, 1976).
86. Kermack, W. O. & McKendrick, A. G. *Proc. R. Soc. A115*, 700–721 (1927).
87. Homer, W. H. *Lancet* **i**, 733–739 (1906).
88. Plowright, R. C. & Paloheimo, J. E. *Theor. Pop. Biol.* **12**, 286–297 (1977).
89. Anderson, R. M. in *Ecological Aspects of Parasitology* (ed. Kennedy, C. R.) (North-Holland, Amsterdam, 1976).
90. Rosenfield, P. L., Smith, R. A. & Wolman, M. G. *Am. J. trop. Med. Hyg.* **26**, 505–516 (1977).
91. Kositzin, V. A. *Symbiose, Parasitisme et Evolution* (Hermann, Paris, 1934).
92. Lewis, T. *Adv. appl. Prob.* **7**, 673–704 (1975).
93. Lotka, A. J. *Am. J. Hyg.* **3**, 1–121 (1923).

articles

Radio studies of the double QSO, 0957+561A, B

G. G. Pooley

Mullard Radio Astronomy Observatory, Cavendish Laboratory, Cambridge, UK

I. W. A. Browne, E. J. Daintree, P. K. Moore, R. G. Noble & D. Walsh

University of Manchester, Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire, UK

The radio source 0957+561 has at least four components. Two coincide with the optical QSOs, which is in accordance with the hypothesis that the QSOs are images of a single object due to a gravitational lens. There are details of spectra and structure which are more difficult to reconcile with the hypothesis.

and Weymann¹ and they suggest that the QSOs may be two images of the same object formed by a gravitational lens. The detailed properties of the radio source are clearly of great interest, and we present here a radio map and other observations discussed in the context of a gravitational lens.

Observations

The total flux density of the radio source has been measured at various frequencies summarised in Table 1. The data are consistent with a flux density $S \propto \nu^{-0.65}$. The two observations at λ 6 cm suggest the possibility of variability.

THE pair of QSOs 0957+561A, B have been shown to have remarkably similar optical characteristics by Walsh, Carswell

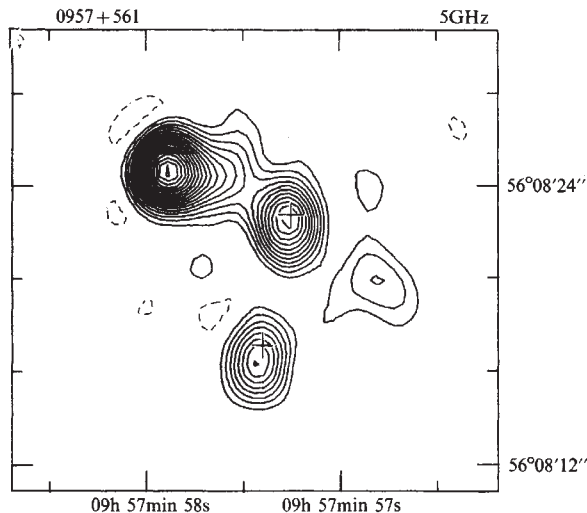


Fig. 1 5-GHz map with 5-km telescope using normal aperture grading. The contour interval is 4 mJy for an unresolved source. Negative contours are shown as dashed lines and the zero contour is omitted. +, Positions of the optical QSOs.

To obtain the 5 GHz map the source was observed with the Cambridge 5-km telescope for three 12-h periods during the period 22–29 May 1979, twice to map the Stokes parameters $I-Q$ and once to map U and Q . The two observations of $I-Q$ show only differences consistent with noise, atmospheric fluctuations, and so on. The mean of the two sets of data is presented in Figs 1 and 2. Figure 1 shows the map produced using the standard truncated gaussian grading function (falling to 0.3 at the largest spacing) with a beamwidth 2" in RA by 2.4" in dec. The map in Fig. 2 was produced using uniform aperture grading with a $1.75'' \times 2.11''$ beam but rather larger sidelobe responses. Note that the declination scale in Fig. 2 is compressed by $\sin \delta$ so that the beam appears circular on this projection; this shows more clearly whether components are extended.

The positions, flux densities and angular sizes of the components are summarised in Table 2, which also includes the positions of the optical objects A and B for comparison.

The polarisation observations show no polarised flux >10 mJy, except for component B, which shows a marginally significant value of 11 mJy at position angle 90° .

Observations were made at 408 MHz in April–May 1979 with interferometers consisting of (1) the MK IA telescope of Jodrell Bank and a 25-m telescope at Knockin (67 km south-west of Jodrell Bank), and (2) MK IA and a 25-m telescope at Defford (127 km south of Jodrell Bank). In each case the source was tracked continuously for 24 h. Data were also taken simultaneously on a compact source, 0954+556 (4C55.17) within the primary beam. The latter was used as a reference to reduce phase fluctuations due to the troposphere and ionosphere using the technique described by Peckham². The data were transformed and 'cleaned' to produce a map with a synthesised beam

of FWHM 0.9". The map suffers limitations because of incomplete coverage of the $u-v$ plane, and residual phase fluctuations. In particular, the incomplete coverage means that components extended on a scale $\geq 4''$ will not be detected.

The maximum correlated flux density is 900 mJy, corresponding to a visibility of ~ 0.6 . Details of the various components definitely detected are given in Table 3. The flux densities are subject to a possible scale error of $\pm 25\%$ and a random error of ~ 50 mJy. At 408 MHz the NE component is the dominant one and its extension is very similar to that seen at 5 GHz. The ratio of the integrated flux densities at these two frequencies has spectral index α_{408}^{5000} of -0.65 , similar to the overall spectrum of the source. It is not possible to say whether component W is extended, but again $\alpha_{408}^{5000} = -0.65$ although the error in this value is rather large. The spectrum of component A appears to be significantly flatter than either NE or W. In addition to the components in Table 3, there is a possible component which may be extended $\sim 2''-3''$, centred east of component B. There is no definite detection of component B itself.

Observations were also made at 1,612 MHz between MK IA and Knockin and at 1,667 MHz between MK IA and Defford, again for 24 h at each baseline. Because of the narrower primary beam it was not possible to observe 0957+561 and the reference 0954+556 simultaneously. Instead each was observed for alternate 10-min intervals throughout the 24 h. The data were combined into a 'clean' map with a synthesised beam of FWHM 0.2", and components extended on a scale $\geq 1''$ cannot be detected. The maximum correlated flux density was 190 mJy corresponding to a visibility of the order ~ 0.4 . Because of the poorer signal-to-noise ratio and the inability to observe the reference source simultaneously, the phase fluctuations were worse than at 408 MHz and the map correspondingly degraded. The apparent flux density of components is reduced by an unknown factor. A compact feature with extension $< 0.2''$ was, nevertheless, clearly detected in component A. There were also probably detections of components W and B; both were apparently extended and this makes comparison of flux densities with A very difficult.

Table 2 Details of components on 5 GHz map; positions of optical QSOs are included for comparison

Component	RA	Dec (1950.0)	Flux density (mJy)	Size (arc s)
W	09 h 57 min 56.81 s	+56° 08' 19.8"	25	1.5–2
A radio	57.27	22.5	48	1
optical	57.25	22.7		
B radio	57.43	16.4	37	1.5
optical	57.40	17.1		
NE	57.90	24.7	120	~ 3

Discussion

The structure of 0957+561 is unusually complex, and unlike any other single radio source we know. We consider first the relevance of the observations to the gravitational lens hypothesis. The most important prediction of the theory is that a point object should give rise to point images which have identical spectra apart from possible extinction along the differing paths of propagation. In the case of a point mass causing the gravitational deflection there should be two images, but if an extended mass such as a galaxy causes the deflection, then multiple images are possible (see ref. 3 and references therein). Radio components A and B in Figs 1 and 2 are coincident with the optical objects within the errors of measurement. Like the optical QSOs they are comparable in flux density. However, a detailed comparison of radio and optical properties presents some apparent discrepancies with the theory.

Flux ratio. Walsh *et al.*¹ pointed out that in the blue the flux density of QSO B is ~ 0.7 that of QSO A but in the red they are

Table 1 Spectral data for 0957+561

Frequency (MHz)	Flux density (mJy)	Error (mJy)	Epoch
151	2,200	200	1978*
408	1,450	150	1979.4†
966	560	150	1976.8§
1,407	450	40	1979.4†
2,695	360	15	1976.7
4,995	230	20	1979.4¶
5,000	310	27	1975.5

Flux densities from the following telescopes: * Cambridge 151 MHz (R. Saunders, personal communication); † Cambridge One-Mile; § Jodrell Bank 278-m interferometer; || NRAO 300-ft (R. W. Porcas, personal communication); ¶ Cambridge 5-km.

equal. They suggested that a possible cause is extinction by dust at the redshift of the QSO ($z = 1.4$). If so then the intrinsic flux from B should be greater than that from A, the ratio being ~ 1.5 . While this factor is somewhat uncertain, it should definitely be > 1 (ref. 4). According to Table 2, the ratio at 5 GHz is ~ 0.8 . If the gravitational lens hypothesis is to accommodate this discrepancy, then (1) the source must vary at either radio or optical wavelengths, or both, or (2) an alternative explanation to the dust hypothesis must be found to account for the optical colour difference. In connection with (1), we note that the two images should follow identical light curves with a relative delay of several months to years depending on the precise geometry of the paths. The two observations at 5 GHz in Table 1 indicate a change in total flux density of 80 ± 35 mJy over 4 yr. While some reservation must be made about this apparent variation because the observations were made by very different techniques, clearly a variation of this magnitude could reconcile the observed flux densities. Changes should be observable in 5 GHz maps fairly quickly, and regular observations are planned for this purpose.

Radio spectral differences. This is an extension of the previous point. The failure to detect component B at 408 MHz suggests the flux ratio (B/A) is ≤ 0.5 , which is less than that at 5 GHz or in the optical region. Again, in principle, this could be accounted for by differences in variation at different frequencies; indeed, variable radio sources do in fact normally show differences in the pattern of variation at different frequencies. However, variation of this magnitude at 408 MHz is very unusual. The errors in the 408 MHz data at present require this point to be treated with caution, but if further observations confirm it then it may be difficult to sustain the gravitational lens hypothesis. We hope to make improved long-baseline observations at 408 MHz shortly.

Extension of the components. Figure 2 suggests that component A is $< 1''$ in extent but B is extended $\sim 1.5''$ along the line joining A and B. This is incompatible with the simple picture of two images of a point object. We may be seeing unresolved multiple images produced because the mass causing the deflection is extended. The direction of extension is as expected for such an effect. The optical image should then also be extended; available observations do not enable this to be seen. Higher resolution observations at both radio and optical wavelengths would clearly be desirable.

So far we have considered only the components A and B close to the optical QSOs. We now consider the components NE and W. Suppose we are in fact seeing images formed by a gravitational lens. Where are the true positions of the 'objects' and where are the second images? Figure 3 shows a possible geometry based purely on the radio observations.

The theory of image formation by a point deflecting mass can be found in, for example, Refsdal⁵ and Press and Gunn⁶. Here we quote some results which are useful in interpreting obser-

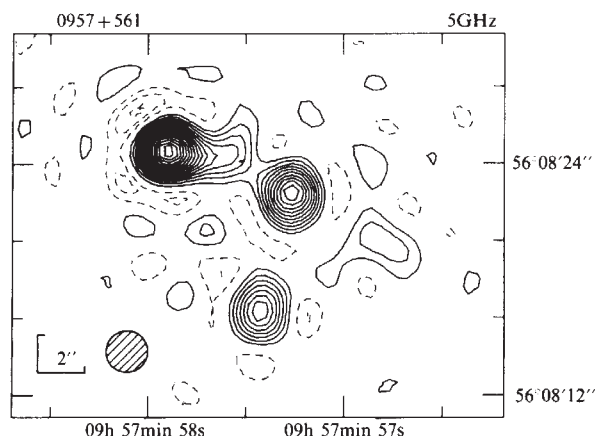


Fig. 2 5-GHz map with uniform aperture grading and declination scale compressed by $\sin \delta$.

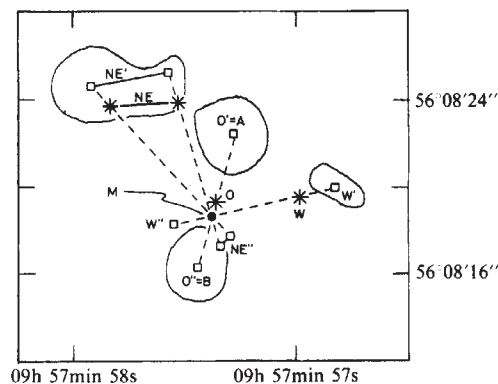


Fig. 3 Possible geometry of the source and image formed by gravitational lens. M is a point deflecting mass; W, O and NE are true positions of 'object' components; W', O' = A, O'' = B and NE', NE'' are the images. The position of M is determined purely from the radio flux density ratio of components A and B.

vations and for constructions such as Fig. 3. On the celestial sphere, let M be the projection of the deflecting mass and O the projection of the 'true' position of a point object (the position in the absence of M). Then point images O', O'' are formed on the extension of the line OM; the former is on the opposite side of O from M, the latter on the opposite side of M from O. Let the angles MO, MO', MO'' and O'O'' be β , α_1 , α_2 and α respectively. Clearly, $\alpha = \alpha_1 + \alpha_2$. Then the following relations hold:

$$\text{Flux ratio of O' and O''} = F'/F'' = \alpha_1^2/\alpha_2^2$$

$$4\alpha_1\alpha_2 = \alpha^2 - \beta^2 = \alpha_0^2 = \text{constant independent of data}$$

$$\alpha_1 = \alpha_2 + \beta$$

The flux ratio and angular separation of A and B with these equations enable Fig. 3 to be constructed.

The second images of components NE and W are in fact too faint to be detected as separate components (less than 1 contour on Fig. 1). Note that they lie in the region of the northwards extension of component B, which may provide an alternative explanation of the extension. The true source then consists of three components: a flat spectrum one giving rise to images A and B (and the optical images) and two steep spectrum ones giving rise to images NE and W. Three-component structure of radio sources is quite common with a flat spectrum object coincident with the optical object and two steep spectrum lobes on opposite sides. In Fig. 3 the true position of the flat spectrum component is not on the line joining the two steep spectrum components. However, though unusual, this configuration is by no means unknown, see for example 3C270.1 and 3C275.1 (ref. 7).

If we are not seeing the result of a gravitational lens, then since component A lies on the line joining NE and W it is likely that these three components constitute a single radio source of a common type. Then we must assume component B is independently a radio source of very nearly the same flux density as A. This then adds another coincidence to those of the emission spectra and absorption spectra described by Walsh *et al.*¹ and by Weymann *et al.*⁸.

Table 3 Details of components detected at 408 MHz

Component	RA	Dec (1950.0)	Flux density (mJy)	Size (arc s)
W	09 h 57 min 56.7 s	+56° 08' 19"	140	—
A	57.36	22.6	120	—
NE	57.89	24.9	600	3

Finally on the gravitational lens hypothesis we consider the possibility that the massive object causing the deflection is itself a radio source. While the mass required depends on the distance of the object, it is likely to be in the range of the most massive elliptical galaxies. Statistics are very sparse for such objects, but using the bivariate luminosity function of Auriemma *et al.*⁹ as a guide, the probability that the galaxy may have an associated radio source of flux density 25 mJy at 5 GHz is ~ 0.005 . Thus it is unlikely that either component W or NE is associated with the lens itself.

Conclusions

The most important feature of the radio data is that there are radio components coincident with the optical QSOs and with comparable flux density at 5 GHz where the structure is best determined. This is consistent with the gravitational lens hypo-

thesis, which would have been ruled out if, for example, only one QSO had a coincident radio source. However, there are details which may or may not be reconciled with the hypothesis. More stringent tests, the most crucial being to study the variations of flux from A and B, must now be made to distinguish between the gravitational lens and other hypotheses.

We thank Tim Cornwell and Peter Wilkinson for assistance with data reduction.

Received 19 June; accepted 13 July 1979.

1. Walsh, D., Carswell, R. F. & Weymann, R. J. *Nature* **279**, 381 (1979).
2. Peckham, R. J. *Mon. Not. R. astr. Soc.* **145**, 31 (1973).
3. Bourassa, R. R. & Kantowski, R. *Astrophys. J.* **205**, 674 (1976).
4. Walsh, D. & Carswell, R. F. (in preparation).
5. Refsdal, S. *Mon. Not. R. astr. Soc.* **128**, 295 (1964).
6. Press, W. H. & Gunn, J. E. *Astrophys. J.* **185**, 397 (1973).
7. Jenkins, C. J., Pooley, G. G. & Riley, J. M. *Mem. R. astr. Soc.* **84**, 61 (1977).
8. Weymann, R. J. *et al. Astrophys. J. Lett.* (in the press).
9. Auriemma, C. *et al. Astr. Astrophys.* **57**, 41 (1977).

The amino acid sequence of the α -chain of human fibrinogen

R. F. Doolittle, K. W. K. Watt, B. A. Cottrell, D. D. Strong & M. Riley

Department of Chemistry, University of California, San Diego, La Jolla, California, 92093

The amino acid sequence of the human fibrinogen α -chain reveals a structure that can be divided into three zones of unique amino acid composition. The middle of these contains the two primary α -chain cross-linking acceptor sites and consists of a remarkable series of internal duplications.

THE α -chain is the largest of the three non-identical chains that compose the human fibrinogen molecule ($\alpha_2\beta_2\gamma_2$)¹. It contains 610 residues and has a calculated molecular weight of 66,125. Unlike the β - and γ -chains, it does not contain covalently bound carbohydrate^{2,3}. More than 90% of the α -chain amino acid sequence has previously been reported⁴⁻¹⁰. In this article we report the complete amino acid sequence (Fig. 1) and comment on a variety of structural and evolutionary features revealed by these determinations. The completion of the sequence was accomplished by the isolation of an 84-residue peptide from a staphylococcal protease digest of the largest α -chain cyanogen bromide fragment. A summary of the data used to obtain the sequence is depicted in Fig. 2.

Distribution of amino acids

A linear residue-by-residue plot of the amino acid sequence of the human fibrinogen α -chain indicates that the structure can be divided into three distinct zones of approximately 200 residues, each of which is readily distinguishable on the basis of amino acid composition alone (Fig. 3). We do not know whether these zones are equivalent to independently folding domains. The regions are designated ZN, ZM and ZC and correspond roughly to the amino-terminal third (ZN), the middle third (ZM), and the carboxy-terminal third (ZC), respectively. A short interzonal sequence between zones ZN and ZM has been designated IZ (Fig. 4). In zone ZN, the first 194 residues from the amino-terminus, the amino acid composition is well balanced except for a very long proline-free stretch and a notable deficiency of glycine residues. This part of the α -chain is linked to the β - and γ -chains by two sets of unique six-cysteine combinations which we have called disulphide rings¹¹. The

regions between these delineating rings are thought to be three-stranded compound α -helices of the sort known as coiled coils¹¹. Indeed, computer-generated conformational prediction schemes¹² indicate that virtually all the α -helical content of the α -chain occurs in zone ZN (Fig. 4).

Zone ZM (middle section) is composed of residues 240-424. More than half the 185 residues in this section are glycine, serine, proline or threonine, and, with the exception of tryptophan, there are few amino acids with nonpolar side chains. The tryptophan exception is extraordinary in that 7 of the α -chain's 10 tryptophan residues occur in this otherwise

```

1  ADSGEGDFLAEGGGVRGPRVVERHQSACKD
31 SDWPFCSDEDWNYKCPSGCRMKGLIDEVNG
61 DFTNRINKLKLNSLFYQKNKDSHSLTTNI
91 MEILRGDFSSANNRONTYNRVSEDLRSRIE
121 VLKRVKVIQKVQHIQLLQKNVRAQLVDMKRL
151 EVDIDIKIRSCRGSCSRALAREVDLKNYED
181 QKKQLEQVIAKDLLPSRDRQHLPLIKMKPV
211 PNLVPGNFKSQLQKVPEWKALTDMPQMRM
241 ELERPGGNEITRGGSTSYGTGSETESPRNP
271 SSAGSWNSGSSGPGSTGNRNPSSSGTGSGA
301 TWKPGSSGPGSTGSWNSGSSGTGSGTGNQNP
331 GSPPRGSTGTWNPSSSERGSAGHMTSESSV
361 SGSTGQWHSESGSFRPDSPGSGNARNPDNP
391 WGTFEVSGNVSPGTRREYHTEKLVTSKGD
421 KELRTGKEKVTSGSTTTTRSCSKVTYTKTV
451 IGPDGHKEVTKEVVTSEDDGSDCPEAMD LGT
481 LSGIGTL DGFRRHPDEAAFFDTASTGKTF
511 PGFFSPMLGEFVSETESRGSESGIFTNTKE
541 SSSHHPGIAEFPSRGKSSSYSKQFTSSTSY
571 NRGDSTFESKSYKMADEAGSEADHEGTHST
601 KRGHAKSRPV

```

Fig. 1 Complete amino acid sequence of the α -chain of human fibrinogen.

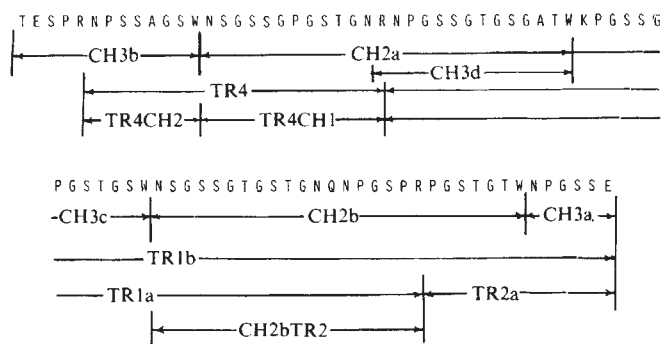


Fig. 2 Summary of data used to deduce the sequence of an 84-residue peptide isolated from a staphylococcal protease digest of the largest α -chain cyanogen bromide fragment. TR = trypsin; CH = chymotrypsin. The sequences of the various subpeptides were determined by procedures reported in our earlier papers. Full experimental details will be presented elsewhere.

exceedingly polar zone. Moreover, the tryptophan residues occur with a definite periodicity reflecting an extended series of 13-residue homology repeats (Fig. 5). The homology repeats, which can also be depicted as a set of 39-residue duplications, have been evaluated by the computer conformation programs and assessed as being predominantly random coil with some contribution from reverse turns. A prototype sequence, derived from a count of the most frequent residues at each position of the 13-residue segments, was also evaluated and found to be a random coil with regularly interspersed turns (Fig. 5).

The carboxy-terminal zone (ZC; residues 425–610) is also rich in serine, threonine and glycine, but in addition, this section contains a large number of amino acids with nonpolar side chains. Although there are no tryptophan residues, more than a quarter of the 186 residues are categorised as nonpolar. A single interchain disulphide loop connects residues 442 and 472. As in the case of ZM, the major conformational prediction is for random coil, although a short segment of α -helix is predicted for the region 583–591.

Distribution of charges

Approximately one-quarter of the α -chain's 610 residues are charged at neutral pH, the net charge being about +6 (Table 1). This net positive charge is partly offset by the fractional occurrence of phosphoserine residues that have not been included in

the estimate. Moreover, each of the major zones is effectively neutral, compensating numbers of positive and negative charges being present throughout. An exception occurs in the 45-residue interzonal segment IZ, where basic residues (neglecting the single histidine) outnumber acidic residues by eight to three. Zone ZM is not only effectively neutral, but it also bears fewer charges generally than the other two major segments; only one-sixth of its side chains are charged at neutral pH (Table 1).

Although we have commented on the general charge distribution of the entire fibrinogen molecule elsewhere¹³, it is useful to summarise some of those observations here (Table 2). Thus, the net charge on the three non-identical chains from fibrinogen as estimated from their amino acid sequences is in complete agreement with their observed behaviour on CM-cellulose chromatography, the most negatively charged γ -chain eluting first, followed by the β -chain, which is less negative, and finally, by the α -chain with its net positive charge. Moreover, the sum of the estimates (Table 2) indicates that at pH 7.3 the parent fibrinogen molecule bears a net charge of about -20. When considered domain by domain, it seems that almost half of this negativity is localised in the central domain (net charge -8) and that a charge reversal occurs centrally on removal of the two pairs of negatively charged fibrinopeptides (A and B) by thrombin. The terminal domains (compare fragments D) are also negative (-4 each), whereas the combined α -chain protuberances (ZM and ZC) are only very slightly negative (-1 each).

Susceptibility to proteases

The susceptibility of the α -chain to proteases—especially as the chain exists in the native fibrinogen molecule—provides certain insights into its domainal structure. In particular, the 45-residue

Table 1 Distribution of charges in α -chain by zone

Zone	Residues	Lys Arg	His	Glu Asp	Approximate net charge*	% Charged residues
ZN	1–194	33	3	32	+2	(34%)
IZ	195–239	8	1	3	+5	(24%)
ZM	240–424	15	3	16	0	(16%)
ZC	425–610	23	8	26	-1	(27%)
Total chain	1–610	79	15	77	(+5)	(26%)

*Calculated for pH 7.3; histidine pK taken as 6.7. The estimate does not include contributions from fractionally occurring phosphoserines.

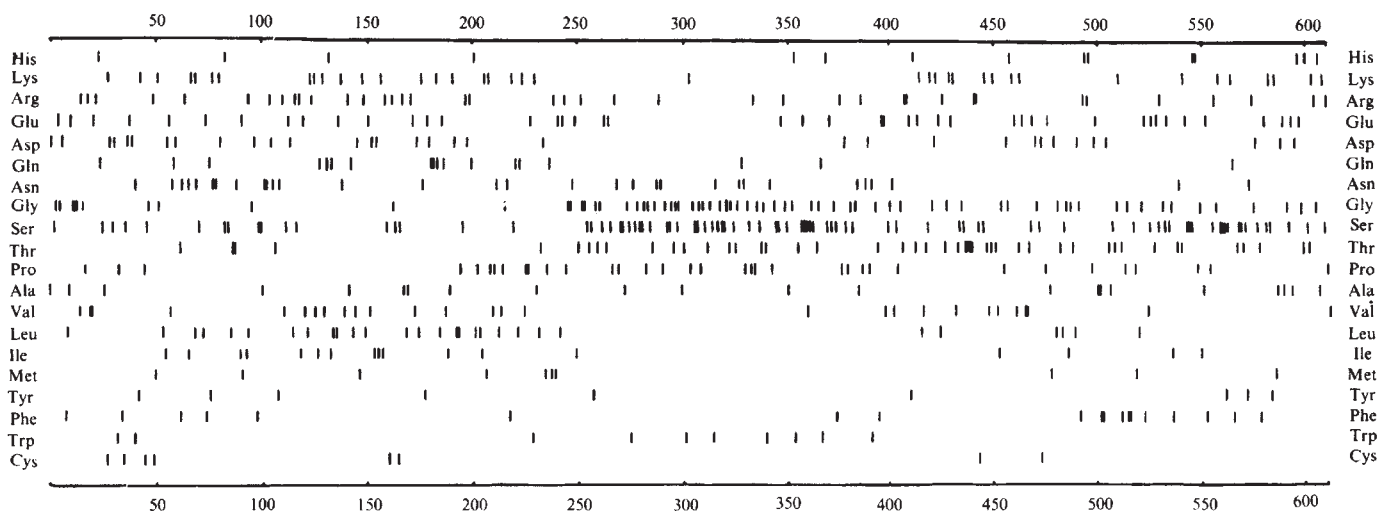


Fig. 3 Distribution of amino acids in the α -chain of human fibrinogen showing how composition changes radically from one end of the chain to the other. Zone N (amino-terminal) includes residues 1–194. The distribution of amino acids is fairly even except for the marked deficiency of proline and glycine. It is separated from the next major sector by a connecting segment (IZ = 195–245) that is rich in proline and exceedingly susceptible to proteolytic excision. Zone M (middle) extends from residues 246 to about 410. More than half the residues in this section are glycine, serine, proline or threonine, but it also contains seven tryptophan residues. The section contains only two glutamyl side chains, both of which are cross-linking acceptor sites. The third sector (ZC) begins at about residue 411 and extends to the carboxy-terminus.

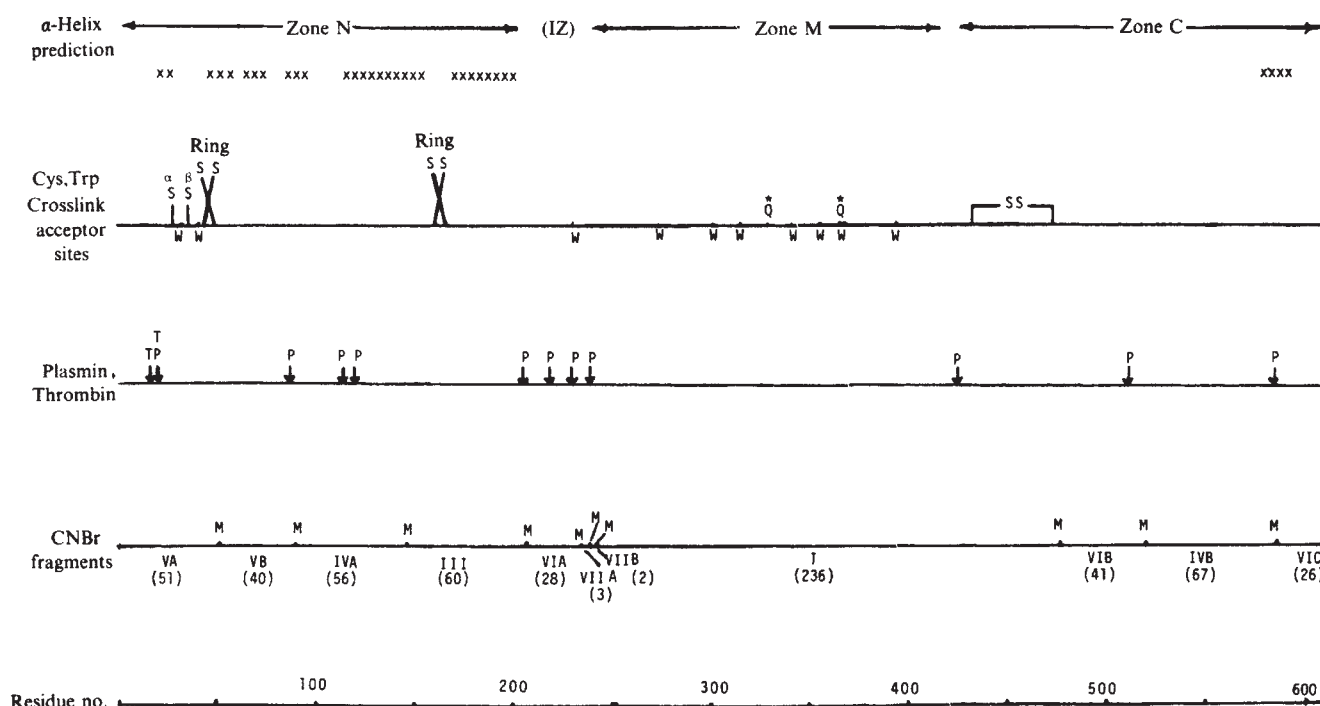


Fig. 4 Schematic representation of human fibrinogen α -chain showing thrombin attack point in releasing fibrinopeptide A(T), principal plasmin attack points (P), cross-linking sites ($\dot{Q}$), 10 tryptophan residues (W) and 8 cysteine residues (C). Also shown are the 11 peptides produced by cyanogen bromide cleavage at the 10 methionine residues (M). Regions of computer-predicted α -helix are denoted by $\times\times\times$.

interzonal (IZ) stretch from residues 195 to 239 includes four plasmin attack points⁵. This short segment contains eight proline residues which reflect the radical change in amino acid composition at this junction and mark the end of the long α -helical section that extends through most of ZN. Its net positive charge (Table 1) may also be a factor in its susceptibility to plasmin. The same region is apparently attacked by several proteases other than plasmin, as evidenced by the fact that digestion of fibrinogen (or fibrin) by a variety of enzymes leads to similar sets of early digestion products, at least on a gross level¹⁴. This kind of substrate-directed vulnerability indicates that the interzonal segment is exposed and accessible.

A second major attack region for plasmin occurs in the middle of the 'coiled-coil' region. The α -chain, in common with the β -

and γ -chains (Fig. 6), is broken at several lysine and arginine residues in this region to give rise (ultimately) to the core fragments known as fragments D and E (ref. 15).

Plasmin also attacks the α -chain and/or its largest cyanogen bromide fragment, CNI (Fig. 4), at a point (Arg 424) separating ZM and ZC. Several other plasmin cleavages also occur, both in the native fibrinogen molecule and in the reduced and alkylated α -chain, especially near the ends of the chain. Finally, as is well known, thrombin triggers the polymerisation of fibrinogen to yield fibrin by cleaving the arginyl-glycine bond at α 16-17, thereby releasing the fibrinopeptide A and exposing the Gly-Pro-Arg sequence at residues α 17-19 which is apparently essential for polymerisation.

Other functional considerations

α -Chain cross-linking: Although the first stages of fibrin polymerisation do not involve covalent bonds, the polymer is subsequently reinforced ('stabilised') by the introduction of intermolecular γ -glutamyl- ϵ -lysine peptide bonds. The formation of these covalent bonds between side chains is catalysed by the transglutaminase known as factor XIII. The first of these bonds formed involves γ -chains exclusively, a set of reciprocal linkages being formed intermolecularly between abutting γ -chains in the polymer, which, on reduction and dissociation of the parent fibrin(ogen) molecules, gives rise to ' γ - γ dimers' (ref.

	1	2	3	4	5	6	7	8	9	0	1	2	3
1	T	E	S	P	R	N	P	S	S	A	G	S	W
2	N	S	G	S	S	G	P	G	S	T	G	N	R
3	N	P	G	S	S	G	T	G	S	G	A	T	W
4	K	P	G	S	S	G	P	G	S	T	G	S	W
5	N	S	G	S	S	G	T	G	S	T	G	N	$\dot{Q}$
6	N	P	G	S	P	R	P	G	S	T	G	T	W
7	N	P	G	S	S	E	R	G	S	A	G	H	W
8	T	S	E	S	S	V	S	G	S	T	G	$\dot{Q}$	W
9	H	S	E	S	G	S	F	R	P	D	S	P	G
10	S	G	N	A	R	P	N	D	P	N	W	.	.
(Prototype)	N	P	G	S	S	G	P	G	S	T	G	T	W
	c	t	c	c	c	c	t	c	c	c	t	c	c

Fig. 5 Amino acid sequence of a section from zone M showing 13-residue homology repeat that occurs 10 times in tandem. A prototype sequence is shown at the bottom together with a computer-derived structural prediction for each residue (c = random coil; t = reverse turn). Asterisks denote cross-link acceptor sites.

Table 2 Distribution of charges on human fibrinogen by chain*

Chain	Lys Arg	His	Glu Asp	Sialic acid	Phospho- serine	Approximate net charge [‡]
α -Chain	79	15	77	0	1 [†]	+5
β -Chain	58	7	61	2	0	-5
γ -Chain	44	10	54	2	0	-10
$\alpha\beta\gamma$	181	32	192	4	1	-10
$\alpha_2\beta_2\gamma_2$	362	64	384	8	2	-20

*Adapted from ref. 13.

[†]Phosphoserine occurs fractionally at two or more locations.

[‡]Calculated for pH 7.3; histidine pK taken as 6.7.

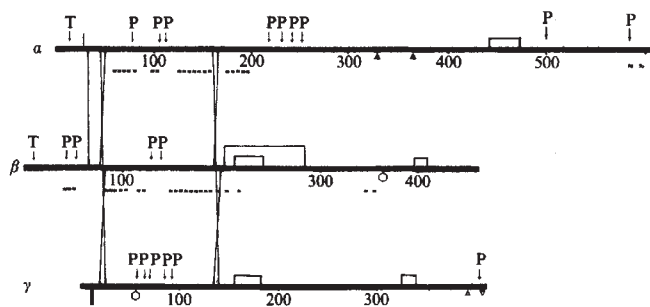


Fig. 6 Schematic depiction of three non-identical chains of human fibrinogen showing disulphide connections, carbohydrate attachment sites (○), thrombin (T) and plasmin (P) attack points, cross-linking sites (Δ and ∇) and computer-predicted helical regions (≈ ≈ ≈).

16). In a significantly slower involvement, α -chains can also be cross-linked by factor XIII, the bonding in this case leading to extended interchain involvements referred to as α -chain multimers¹⁷. However, the overall geometry and arrangement of these multimeric arrays has been unclear. We can now report the exact position of those α -chain glutamine residues that incorporate radioactive substitute donors (¹⁴C-glycine ethyl-ester) under the influence of activated factor XIII. Thus, glutamines α 328 and α 366, both of which occur in ZM, are the only acceptor sites for the substitute donor. Significantly, however, they are also the only glutamine residues in ZM, thus raising the question of whether the specificity is narrow because there are only two glutamine residues or vice versa, an evolutionary exclusion of glutamine residues being required to maintain the ordered cross-linking of these randomly disposed segments. Furthermore, there is only a single glutamine

residue in the carboxyl-terminal zone ZC (Gln 563), and it was devoid of radioactivity; nor were any of the 15 glutamine residues in ZN and IZ labelled. Furthermore, in neither Gln 328 nor Gln 366 does the surrounding sequence bear any resemblance to the γ -chain cross-linking acceptor site¹⁶.

Other α -chain involvements: Fibrinogen is widely known to interact with certain other plasma proteins and particles, including associations with fibronectin ('cold-insoluble globulin')^{18,19}, platelets²⁰ and certain strains of staphylococcal bacteria²¹. These interactions are thought to involve the carboxyl two-thirds of α -chains, evidence for this including the observation that plasmin digestion of fibrinogen removes those regions concomitant with a loss of interaction²²⁻²⁴.

During our determination of the α -chain structure we developed a procedure for isolating a soluble fragment corresponding to residues 241-424 ($n = 184$ residues). This fragment, which is a plasmin digestion product of the cyanogen bromide fragment CNI, includes both cross-linking acceptor sites. It also contains two tyrosine residues which can be iodinated. Thus, the fragment should prove very useful in pinpointing many of the extramolecular involvements cited above.

The α -chain protuberance comprised of ZM and ZC also affects other aspects of fibrinogen-fibrin behaviour. It is well known, for example, that fibrinogen 'sticks' to most non-biological materials, including the plastics used in most circulatory-assistance devices. It is possible that this is the result of exposing the tryptophan residues that are normally 'caged' by the repetitive Pro-Gly-Ser-Thr combinations. Also, these portions of the protein are very accessible to cross-linking agents and other types of chemical derivatisation²⁵. Furthermore, it has so far not been possible to crystallise mammalian fibrinogens until portions of the α -chain have been digested away²⁶. Finally, high-solubility fibrinogen preparations seem to lack these regions of the α -chain.²⁷

α	ADSGEGDFLAEGGGVVRGPRVVERHQSA	CKDSWPFC	SDEDWNYKCPSGCRMKGILDEVNQDFTNRINKLNKSLFEYQKNK	DSHSLTT
β	QGVNDNEEGFFSARGHRLDKKREAPSLRPAPPISGGGYRARPAAATQKKVERKAPDAGGCLHADPDGLVLCPTGCQLQEALLQQRPIRNSVDELNNNVEAVSQTSSSQFYMYL			
γ			YVATRDNCILDERFGSYCPTTCGIADFLSTYQTKVDKDLQSLDILHQVENKTS	EVKQLIK
α	NIMEILRGDFSSANNRNTYNRVSEDLRSRIEVLKRKVIQKVQHIQLLEKNVRAQLVDMKRLVDIDIKIRSCRGSCSRALAREVDLKNYEDQKQLEQVIAKDLLPSRDRQHLPLIKMK			
β	LKDLWQKRQKQVDNENNVNYSSELEKHQLYIDETVNSNIPTNLRLVLSILENLRSKIQLKESDVSQAQMEYCRTPCTVSCDIPVVS		KECEEIIRKGGTSEMYLIQPDSSVK	
γ	AIQLTYNPDESSKPNMIDAATLKSRLKEEIMKYEASILTHDSSIRYLQEIYNSNNQKIVNLKEKVAQLEAQCEPCDKDTVQIHDTG		KDCQDIANKGAKQSGLYFIKPLKANQ	
α	PVPNLVPGNFKSQLKQVPPPEWKALTDMPQMRM	ELERPGGNEITRGGSTSYGTGSETESPRNPSSAGSWNSGSSGPGSTGNRNPSSGSGTGATW		
β	PYRVYCDMNTENGWTVIQNRQDGSVDGFRKWDPKYQGFQGNV	ATNDGKBKYGCLPGEYWLGBBKISELTRMGPTLLI	EMEDWKGDKYKAHYGGFTYQNEANKYQISVKNKYRGTAGNA	
γ	QFLVYCEIDGSGNGMTVFQKRLDGSVDGFKMWIQYKEGFHLSPTGTT	EFWLGNKIHILISTQSAIPYALRVELEDWNGRTSTADYAMFKYGPEDAKYRLTYAYFAGGDAGD		
α	KPGSSGPGSTGSWNSGSSGTSTGNQHPGSPRPGSTGTWNPSSERGSAGHWTSESSVSGSTGQWHSSEGSFRPDSPGSGNARNPDNMGTFEEVSGNVSPGTRREYHTEKLVTSKGDKELR			
β			LMDGASQL	MGENRTMIHNGMFFS
γ			AFDGFDFGDDPSDKFFTSHNGMQFS	
α	TGKEKVTSGSTTTTTRSCSKTVTKTVIGPDGHKEVTKEYVTSDEGSDCEAMDGLTSLGIGTLDFGRHHRPDAAFFDTASTGKTFPGFFSPMLGEFVSETESRGSESGIFTNTKESSSHHP			
β	TYDRDNDGWLTS DPRKQCSKEDGGGWWYBRCHAANPNRGYYNGGQYTDMAKHGTBDGVVMMNWKGSWYSRKMMS		KIRPFFPQQ	
γ	TWDNDNDKFEQN	CAEQDQSGWMMNKCHAGHLNGVYQGGTYSKASTPNGYDNGIIVAT	KTRWYSMKKTTM	KIIPFNRLTIGEGQHHLLGGAKQAGDV
α	GIAEFPSRGKSSYSKQFTSSTSYNRGDSTFESKYSYMADEAGSEADHEGHTSTKRGHAKSRPV			
β				
γ				

Fig. 7 Comparison of α -chain amino acid sequence with sequences of β - and γ -chains³¹⁻³⁷ showing significant homology among all three chains over the course of the first 210 residues, but a loss of homology in the α -chain thereafter. The 'best match' leaves a gap throughout most of zone M, some faint resemblance between the α -chain and the other two chains apparently resuming near the carboxyl-terminal segment. Other equivalent alignments are possible, however. As previously noted, the β - and γ -chains are strongly homologous over the course of their carboxy-terminal two-thirds. Matched identical residues are underlined.

Evolutionary significance

The α -, β - and γ -chains of fibrinogen are all equivalently homologous in the region spanned by ZN²⁸. Thus, all three chains have two characteristic Cys-X-X-X-Cys sequences spaced appropriately for participation in the two sets of disulphide rings (Fig. 6). Also, both the α - and β -chains have thrombin-sensitive sites near their amino-termini that give rise to the fibrinopeptides A and B, respectively (Fig. 6). Shortly beyond the beginning of ZM, however, the resemblance between the α -chain and the β - and γ -chains abruptly disappears. In contrast, beyond that point the β - and the γ -chains display increasing homology^{29,30}. Figure 7 shows a computer-assisted alignment of the three chains. Other reasonable alignments are possible, however, including one in which the carboxyl-terminal segments of the α - and γ -chains are directly aligned.

Received 26 April; accepted 25 June 1979.

- McKee, P. A., Rogers, L. A., Marler, E. & Hill, R. L. *Archs Biochem. Biophys.* **116**, 271–279 (1966).
- Gaffney, P. J. *Biochim. biophys. Acta* **263**, 453–458 (1972).
- Pizzo, S. V., Schwartz, M. L., Hill, R. L. & McKee, P. A. *J. biol. Chem.* **247**, 636–645 (1972).
- Blomback, B., Hessel, B., Iwanaga, S., Reuterby, J. & Blomback, M. *J. biol. Chem.* **247**, 1496–1512 (1972).
- Takagi, T. & Doolittle, R. F. *Biochemistry* **14**, 5146–5149 (1975); *Thromb. Res.* **7**, 813–818 (1975).
- Gårdlund, B. *Thromb. Res.* **10**, 689–702 (1977).
- Doolittle, R. F. *et al. Biochemistry* **16**, 1703–1709, 1710–1715 (1977).
- Cottrell, B. A. & Doolittle, R. F. *Biochem. biophys. Res. Commun.* **71**, 745–761 (1976); *Thromb. Res.* **12**, 1135–1146 (1978).
- Lottspeich, F. & Henschen, A. *Hoppe-Seyler's Z. physiol. Chem.* **359**, 1451–1455, 1611–1616 (1978).
- Doolittle, R. F., Cottrell, B. A., Strong, D. & Watt, K. W. K. *Thromb. Res.* **14**, 787–792 (1979).
- Doolittle, R. F., Goldbaum, D. M. & Doolittle, L. R. *J. molec. Biol.* **120**, 311–325 (1978).
- Garnier, J., Osguthorpe, D. J. & Robson, B. *J. molec. Biol.* **120**, 97–120 (1978).
- Doolittle, R. F. *et al. in Chemistry and Physiology of Plasma Proteins* (Pergamon, New York, in the press).
- Gaffney, P. J. in *Haemostasis: Biochemistry, Physiology and Pathology* (eds Ogston, D. & Bennett, B.) 105–168 (John Wiley, London, 1977).
- Nussenzweig, V. *et al. Annls Inst. Pasteur, Paris* **100**, 377–389 (1961).

Given that the prototype fibrinogen molecule originated from two pairs of three identical chains, the question arises, did the β - and γ -chains lose the equivalent of zone ZM, or did it evolve in the α -chain by a more recent set of tandem duplication? In either event, the strong homology between the β - and γ -chain carboxyl-terminal segments suggests that these two chains shared a common ancestor more recently than either did with the α -chain. Moreover, the radical differences in the nature of the various α -chain zones suggests the occurrence of events other than simple gene duplication in the evolution of the α -chain. It is possible that gene splicing was involved, although we know of no other protein with a composition comparable to ZM that might be a candidate for insertion.

The International Committee on Blood Coagulation Nomenclature prefers the convention A α ₂B β ₂ γ ₂ as a description of the six chains in the fibrinogen molecule.

- Chen, R. & Doolittle, R. F. *Proc. natn. Acad. Sci. U.S.A.* **63**, 420–427 (1969); **66**, 472–479 (1970); *Biochemistry* **10**, 4486–4491 (1971).
- McKee, P. A., Mattock, P. & Hill, R. L. *Proc. natn. Acad. Sci. U.S.A.* **66**, 738–744 (1970).
- Mosher, D. F. *J. biol. Chem.* **250**, 6614–6621 (1975).
- Ruoslathi, E. & Vaheri, A. *J. exp. Med.* **141**, 497–501 (1975).
- Brinkhous, K. W., Read, M. S. & Mason, R. G. *Lab. Invest.* **14**, 335–342 (1965).
- Duthie, E. S. *J. gen. Microbiol.* **10**, 427–436 (1954).
- Budzynski, A. Z., Joseph, R. T., Olexa, S. A. & Niewiarowski, S. *Fedn Proc.* **38**, 996 abstr. (1979).
- Hawinger, J. *et al. (in preparation)*.
- Mosesson, M. W. *Ann. N.Y. Acad. Sci.* **312**, 11–30 (1979).
- Doolittle, R. F. *Adv. Protein Chem.* **27**, 1–109 (1973).
- Weisel, J. W., Warren, S. G. & Cohen, C. *J. molec. Biol.* **126**, 159–183 (1978).
- Mosesson, M. W., Umfleet, R. A., Finlayson, J. S. & Galanakis, D. K. *J. biol. Chem.* **247**, 5210–5219 (1972).
- Doolittle, R. F. *Fedn Proc.* **35**, 2145–2149 (1976).
- Henschen, A. & Lottspeich, F. *Thromb. Res.* **11**, 869–880 (1977).
- Watt, K. W. K., Takagi, T. & Doolittle, R. F. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1731–1735 (1978).
- Blomback, B. *et al. J. biol. Chem.* **248**, 5806–5820 (1973).
- Sharp, J. J., Cassman, K. G., & Doolittle, R. F. *FEBS Lett.* **24**, 334–336 (1972).
- Takagi, T. & Doolittle, R. F. *Biochemistry* **14**, 940–946 (1975).
- Henschen, A. & Lottspeich, F. *Hoppe-Seyler's Z. physiol. Chem.* **358**, 935–938 (1977).
- Blomback, B., Hessel, B. & Hogg, D. *Thromb. Res.* **8**, 639–658 (1976).
- Henschen, A. & Lottspeich, F. *Hoppe-Seyler's Z. physiol. Chem.* **358**, 1643–1646 (1977).
- Watt, K. W. K., Takagi, T. & Doolittle, R. F. *Biochemistry* **18**, 68–76 (1979).

The membrane attachment protein for spectrin is associated with band 3 in human erythrocyte membranes

Vann Bennett & Peter J. Stenbuck

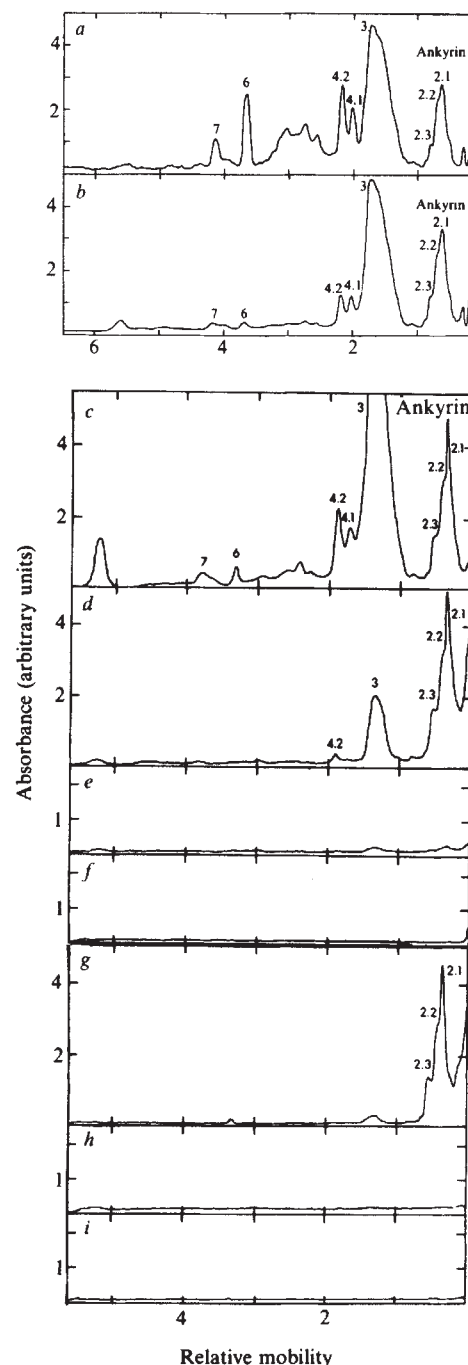
Department of Molecular Biology, The Wellcome Research Laboratories, Research Triangle Park, North Carolina 27709

Ankyrin, the membrane attachment protein for human erythrocyte spectrin, is tightly linked in a 1:1 molar ratio with band 3 in detergent extracts of spectrin-depleted membranes. Ankyrin-linked band 3, which represents 10–15% of the total band 3, spans the membrane, and is nearly identical to the major band 3 by peptide analysis. Spectrin binds to solubilised ankyrin-linked band 3, but not to free band 3. A portion of band 3 remains firmly associated with detergent-extracted cytoskeletal proteins. It is concluded that a fraction of band 3 is attached to the erythrocyte cytoskeleton through association with ankyrin, which in turn is bound to spectrin.

may be key events in complex cellular phenomena such as the capping of surface molecules by lymphocytes⁵, redistribution and internalisation of receptors for hormones⁶ and low-density lipoprotein⁷, cell motility⁸ and attachment of cells to surfaces⁹. Evidence has been presented for binding of actin^{10,11} and intermediate filament protein¹² to cell-surface proteins. However, there is little detailed knowledge of the polypeptides involved in these linkages or of the mechanisms which regulate such interactions. The human erythrocyte provides an experimentally accessible model system for examining protein-membrane interactions at a molecular level. The shape and elasticity of red cell membranes^{13,14} and the distribution of intramembrane particles¹⁵ and glycoproteins¹⁶ are controlled by proteins arranged in a cytoskeletal network on the inner surface of the membrane. Direct analysis of the components involved in these phenomena can be achieved, in principle, by measuring reassociation of purified cytoplasmic proteins with acceptor sites exposed on inverted membrane vesicles. This approach has been used to examine the membrane attachment site for human erythrocyte spectrin, which is the major protein in the red cell cytoskeleton.

EXTENSIONS of the fluid mosaic model for membranes^{1,2} have proposed a direct association between membrane-spanning cell-surface proteins and cytoplasmic structural proteins on the inner surface of the plasma membrane^{3,4}. These protein interactions

Fig. 1 Isolation of ankyrin-associated proteins by immunoadsorption with anti-ankyrin IgG of Triton X-100-solubilised spectrin-depleted vesicles labelled with [125 I]-Bolton Hunter reagent. Spectrin-depleted inverted vesicles¹⁸ from 1-ml erythrocyte ghosts were reacted with 2 mCi [125 I]-Bolton Hunter reagent (1,500 Ci mmol⁻¹, NEN) in 0.2 ml 7.5 mM sodium phosphate, pH 7.5, for 60 min at 0°C. The labelled vesicles (1.1 mCi incorporated) were washed, extracted (15 min, 24°C) with 1.2 ml 7.5 mM sodium phosphate, pH 7.5, 1 mM NaEDTA, 0.2 mM dithiothreitol, 10 μ g ml⁻¹ phenylmethylsulphonyl fluoride and 2% (v/v) Triton X-100, and pelleted (30 min, 250,000g at 2°C). The supernatant was made 0.1 M in KCl, and 10- μ l aliquots were either denatured with an equal volume of 1% (w/v) SDS (30 min, 37°C) or untreated. These samples were incubated (90 min, 0°C) in 0.3 ml buffer containing 250 μ g ml⁻¹ bovine serum albumin, 2.5 μ g ml⁻¹ pancreatic trypsin inhibitor, 100 mM NaCl, 10 mM sodium phosphate, pH 7.5, 1 mM NaEDTA, 0.2% Triton X-100, 0.25 μ l protein A-bearing formalin-fixed staphylococci (Calbiochem) and either 2.5 μ g affinity-purified anti-ankyrin IgG²⁰ or 2.5 μ g preimmune IgG, and in two tubes 15 μ g of the spectrin-binding fragment of ankyrin¹⁹. The reaction was stopped by addition of 3.5 ml 2 M urea, 1% Triton X-100 and 0.1 M glycine, and the bacteria collected by centrifugation (15 min, 4,000g). The bacteria with adsorbed immune complexes were resuspended with 0.1 ml 1% SDS, 40 mM dithiothreitol, 10 mM Tris HCl, pH 8, and 1 mM NaEDTA, and 20- μ l portions analysed by SDS electrophoresis on a 1.5-mm thick 7.5% polyacrylamide slab gel⁴⁹. The gel was stained with Coomassie blue, dried and developed by autoradiography (bottom panel). The detergent extract used for immunoadsorption was diluted to an equivalent volume to that of the precipitates for analysis. Spectrin-depleted everted vesicles used in this experiment were also electrophoresed, and the gels stained with Coomassie blue and visualised by autoradiography. *a*, Coomassie blue stained; *b* [125 I]-labelled; *c*, Triton X-100 extract; *d*, immunoprecipitated extract + anti-ankyrin; *e*, extract + ankyrin fragment + anti-ankyrin; *f*, extract + pre-immune IgG; *g*, immunoprecipitated SDS-denatured extract + anti-ankyrin; *h*, SDS-denatured extract + ankyrin fragment + anti-ankyrin; *i*, SDS-denatured extract + preimmune IgG.



Spectrin has been radiolabelled with 32 P_i in intact erythrocytes and purified as an 8S dimer which can recombine with inverted vesicles depleted of spectrin and actin¹⁷. 32 P-Spectrin binds to vesicles with a K_D of 10^{-7} – 10^{-8} M, binding to a single class of protein sites tightly bound to the inner surface of the membrane, and present in about 10^5 copies per cell¹⁸. A water-soluble, 72,000-molecular weight (MW) fragment of the attachment protein is released by mild proteolysis of inverted vesicles, and this polypeptide has been isolated and characterised¹⁹. The purified attachment fragment binds to spectrin with a 1:1 molar ratio, and competes for reassociation of spectrin with vesicles¹⁹. A 200,000-MW protein localised on the cytoplasmic membrane surface has been identified as the precursor of the spectrin-binding fragment, on the basis of cross-reactivity with mono-specific anti-fragment IgG (ref. 20) and by peptide mapping²¹. This protein, which has been named ankyrin, has been partially purified and found to bind to spectrin as well as competitively inhibiting reassociation of spectrin with the membrane²⁰.

Ankyrin dissociates from the membrane following treatment with dilute acid²⁰, and thus is not an integral membrane protein. It probably does interact with such proteins, as the lateral mobility of fluorescence-labelled integral proteins is increased by incubation of erythrocyte ghosts with the spectrin-binding fragment of ankyrin²². This report demonstrates association of ankyrin with band 3, the major membrane-spanning polypeptide of the erythrocyte membrane^{23–25}.

Identification of ankyrin-associated polypeptides

Ankyrin-associated polypeptides were isolated from Triton X-100-solubilised, spectrin-depleted vesicles by immunoadsorption with monospecific anti-ankyrin IgG. The antiserum was raised against the 72,000-MW spectrin-binding fragment of ankyrin which had been isolated by SDS-polyacrylamide gel electrophoresis (PAGE), and ankyrin-specific IgG was purified by affinity chromatography with immobilised fragment²⁰. The immune complexes formed between anti-ankyrin and the membrane proteins were isolated with formalin-fixed protein

A-bearing staphylococci, which bind with high affinity to the Fc portion of IgG²⁶. The bacteria were subsequently washed with 2 M urea and 1% Triton X-100 to minimise nonspecific interactions. It should be emphasised that only proteins that are tightly associated and unaffected by Triton X-100 will be detected by this procedure.

Membrane polypeptides were analysed in nanogramme amounts without interference from immunoglobulin chains by labelling vesicles with [125 I]-labelled Bolton Hunter reagent. This reaction is achieved in mild conditions, and radiolabels identical polypeptides as visualised by staining with Coomassie blue. Band 3 and ankyrin (bands 2.1, 2.2, 2.3) are labelled in nearly the same relative intensity, although other polypeptides are labelled to a lesser extent by the Bolton Hunter reagent (Fig. 1).

Immunoadsorption of solubilised vesicles with anti-ankyrin IgG resulted in binding of essentially all the 200,000-MW polypeptide (ankyrin) and of polypeptides of ~180,000 and 190,000 MW (bands 2.3 and 2.2) which are proteolytic fragments of ankyrin^{20,27}. In addition, the immunoprecipitate

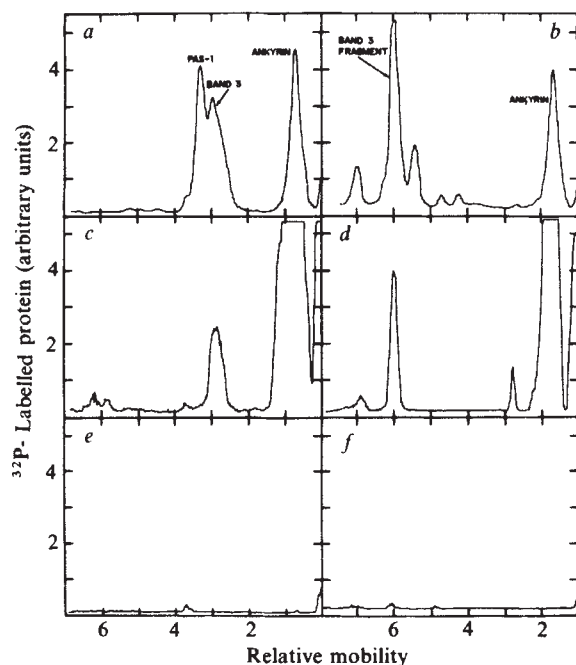


Fig. 2 Immunoadsorption with anti-ankyrin IgG of Triton X-100-solubilised spectrin-depleted vesicles from ^{32}P -labelled erythrocytes that were untreated or externally digested with α -chymotrypsin. Packed erythrocytes (3 ml) were incubated for 150 min at 37°C in the presence or absence of $100\ \mu\text{g ml}^{-1}$ α -chymotrypsin (70 U per mg) in 20 ml of Krebs-Ringer bicarbonate buffer supplemented with $200\ \text{U ml}^{-1}$ penicillin G and 10 mM D-glucose. The cells were washed three times in 0.15 M NaCl, and then labelled with ^{32}P during a 15-h incubation at 37°C in 5 ml of buffer containing 135 mM NaCl, 20 mM NaHCO_3 , 4 mM KCl, 2.5 mM MgCl_2 , 1 mM CaCl_2 , 10 mM D-glucose, 1 mM adenosine, $200\ \text{U ml}^{-1}$ penicillin G, $50\ \mu\text{g ml}^{-1}$ phenylmethylsulphonyl fluoride and 0.5 mM ^{32}P NaPO₄ (4 Ci mmol⁻¹, NEN), pH 7.5. Spectrin-depleted inverted vesicles were prepared from the ^{32}P -labelled cells¹⁸, which were solubilised, immunoadsorbed and analysed as described in Fig. 1. The data are presented as densitometer scans of the portions of autoradiograms corresponding to MWs greater than 50,000. Band 3 and band 3 fragment were identified by comparison with the Coomassie blue-stained gels. *a*, Undigested ^{32}P -labelled vesicles, Triton X-100 extract; *b*, externally digested ^{32}P -labelled vesicles, Triton X-100 extract; *c*, immunoprecipitate of undigested extract + anti-ankyrin; *d*, immunoprecipitate of externally digested extract + anti-ankyrin; *e*, immunoprecipitate of undigested extract + ankyrin fragment + anti-ankyrin; *f*, immunoprecipitate of externally digested extract + ankyrin fragment + anti-ankyrin.

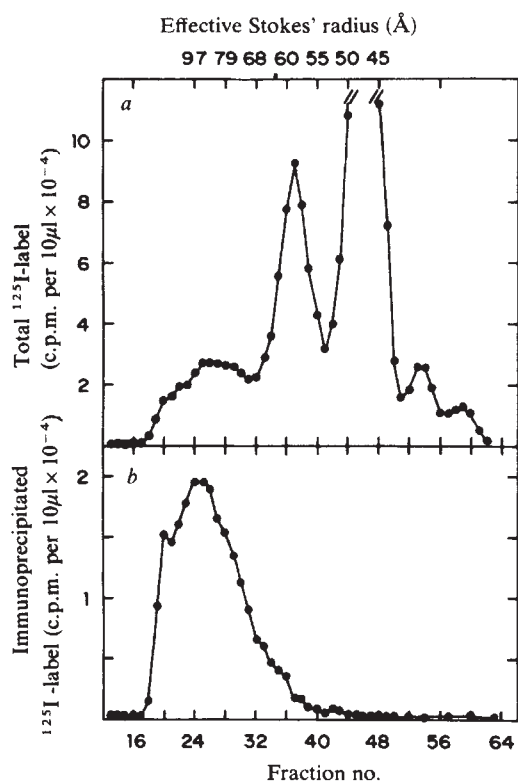


Fig. 3 Ultrogel AcA22 chromatography of Triton X-100-solubilised spectrin-depleted vesicles labelled with ^{125}I -Bolton Hunter reagent. Vesicles were labelled and solubilised (Fig. 1) and a 0.7-ml sample chromatographed at 4°C on a column ($1.0 \times 60\ \text{cm}$) packed with Ultrogel AcA22 (LKB) and equilibrated with 100 mM KCl, 10 mM NaPO_4 , pH 7.5, 1 mM NaEDTA, 1 mM NaN_3 , 0.2 mM dithiothreitol and 0.2% Triton X-100. Fractions (0.8 ml) were assayed for total radioactivity (*a*) and for immunoprecipitable radioactivity using anti-ankyrin fragment IgG as in Fig. 1 (*b*). Pooled fractions (*c*) and the corresponding immunoprecipitates (*d*) were also analysed by SDS electrophoresis. The column was calibrated with rabbit muscle glyceraldehyde-3-phosphate dehydrogenase (44 Å), bovine liver catalase (52 Å), horse spleen ferritin (59 Å) and human IgM (120 Å).

exhibited a diffusely staining band of ~95,000 MW, which contains about 15% of the band 3 material, and a 72,000-MW polypeptide (band 4.2) (Fig. 1). Polypeptides of 28,000 MW (band 7) and 17,000 MW (band 8) were also present, but in smaller amounts which are not well visualised in these experiments. Glycophorin (PAS-1) was not clearly labelled, but was apparent in membranes from erythrocytes incubated with ^{32}P (ref. 28) (Fig. 2), and following periodate oxidation and tritiation with $[^3\text{H}]\text{NaBH}_4$. ^{32}P -Glycophorin was not immunoprecipitated (Fig. 2), and the same result was obtained with the tritium-labelled protein (not shown).

Several lines of evidence demonstrate that the polypeptides immunoprecipitated by anti-ankyrin IgG were associated with ankyrin, and were not nonspecifically adsorbed or bound directly to the antibody. Unlabelled ankyrin fragment blocked binding of ankyrin as well as of band 3 and the other polypeptides (Fig. 1). Immunoabsorption of solubilised vesicles previously denatured with SDS yielded only ankyrin (bands 2.1, 2.2, 2.3), with no band 3 or band 4.2, and this interaction was also blocked by the ankyrin fragment (Fig. 1). Preimmune IgG precipitated no polypeptides from either denatured or untreated vesicle extracts (Fig. 1). Finally, bands 3 and 4.2, which had been separated from ankyrin by gel filtration (see below), did not cross-react with anti-ankyrin IgG.

Ankyrin-associated polypeptides dissociated very slowly from the ankyrin-IgG complex, with less than 10% loss during a 90-min incubation at 37 °C (not shown). Conditions have not yet been developed to separate ankyrin from these proteins without using denaturing conditions. It was therefore not feasible to carry out reassociation experiments with isolated components of the complex.

Gel filtration of solubilised membranes on an Ultrogel AcA22 column separated ankyrin-linked band 3 from the major pool of band 3 in mild conditions (Fig. 4) and confirms the existence of an oligomer containing ankyrin, and bands 3 and 4.2 as its principal components. Immunoabsorption of the fractions with anti-ankyrin IgG precipitated a single broad peak which migrated with an effective Stokes' radius of about 90 Å, and preceded the major peaks of radioactivity (Fig. 3). Ankyrin chromatographed coincidentally with immunoreactive material, and was associated with bands 3 and 4.2. The relative proportions of ankyrin, band 3 and band 4.2 were about the same in the peak fractions as in the immunoprecipitates, and nearly all the radioactivity was immunoreactive. These fractions re-chromatograph at about the same position (not shown). The major peak of band 3 migrated with band 4.2 with a Stokes' radius of about 58 Å, and was not adsorbed by anti-ankyrin IgG (Fig. 3). Glycophorin labelled with ^{32}P or tritium migrated more slowly than band 3. These results agree with the previous conclusion that bands 3 and 4.2 are associated in detergent extracts, whereas band 3 and glycophorin are dissociated in these conditions³⁰. Approximately 80% of band 3 migrated separately from ankyrin, in agreement with the amount of band 3 immunoabsorbed in whole extracts (Fig. 1).

Immunoabsorbed band 3 was present in a 1:1 molar ratio with ankyrin (that is, bands 2.1, 2.2, 2.3) (Fig. 1). (The stoichiometry of band 3 and ankyrin was determined by integration of the area under the respective peaks on autoradiograms, assuming average MWs of 95,000 for band 3 and 200,000 for ankyrin, and that the specific activity of the iodinated polypeptides was the same. This latter assumption is supported by fact that band 3 and ankyrin are labelled in the same relative proportions by Coomassie blue staining as by iodination (Fig. 1).) The major band 3 protein exists as a dimer in erythrocyte ghosts and in Triton X-100-solubilised extracts³⁰⁻³³. Band 3 (B3) and ankyrin (A) could thus be arranged as $((\text{B3})_2-(\text{A})_2)_N$ if the band 3 association remained intact during isolation, or as $((\text{B3})_2-\text{A})_N$ if the band 3 dimer dissociated in these conditions. Ankyrin and band 3 co-migrated on gel filtration in about a 1:1 molar ratio in mild conditions in which band 3 probably exists as a dimer³⁰. It is therefore likely that the ankyrin-linked band 3 oligomer is arranged with 2 mol each of band 3 and ankyrin. This

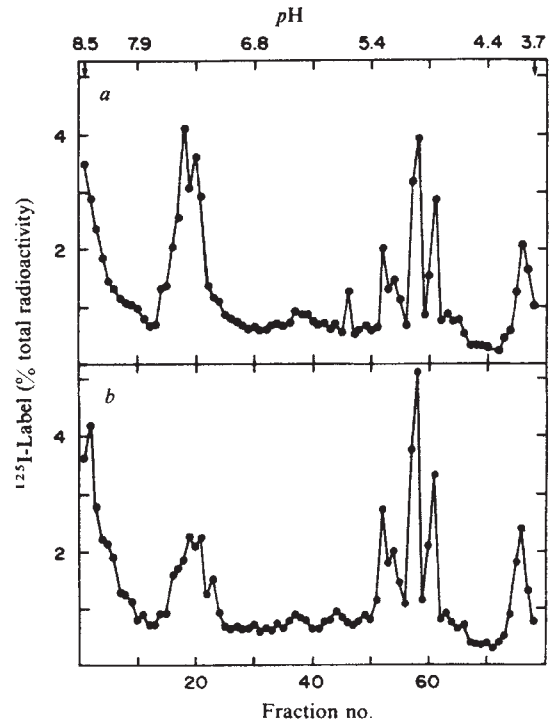


Fig. 4 Isoelectric focusing in polyacrylamide of cyanogen bromide fragments of [^{125}I]-Bolton Hunter-labelled ankyrin-linked band 3 (a), and free band 3 (b). Ankyrin-linked band 3, isolated by immunoabsorption of solubilised [^{125}I]-labelled vesicles (Fig. 1), and free band 3 isolated by gel filtration (Fig. 3) were purified by SDS electrophoresis (Fig. 1). The bands (2×10^5 c.p.m.) were stained with Coomassie blue, cut out and incubated (5 h, 50 °C) in 1 ml of 70% (v/v) formic acid with 40 mg cyanogen bromide. The solutions were separated from polyacrylamide, dried and incubated with 1 M NH_4OH to convert homoserine lactone to the free acid. The samples were reconstituted with 50 μl 9 M urea, 20 mM dithiothreitol, 2% (w/v) ampholines, pH 3.5–10 (LKB), and 20 $\mu\text{g ml}^{-1}$ bromophenol blue and analysed by isoelectric focusing (12 h, 1 W) on a 12.5% 1.5-mm polyacrylamide slab gel containing 9 M urea, 1.3% pH 3.5–10, 1% pH 5–8, 1.3% pH 4–6 ampholines and 0.2% Triton X-100. The gels were sliced and radioactivity determined. The data are expressed as the per cent of the total radioactivity recovered from the gel.

conclusion is supported by the finding that band 2.1 is cross-linked in a dimeric form³¹. In some gel filtration experiments the ratio of band 3 to ankyrin was greater than that of Fig. 3. Thus, additional associations may occur between the ankyrin-band 3–4.2 oligomer and other band 3 molecules, although experimental conditions for stabilising these complexes have not yet been developed.

Ankyrin-linked band 3 is indistinguishable from free band 3

It was important to determine whether ankyrin-associated band 3 has similar properties to those attributed to the major band 3 protein. Ankyrin-linked band 3 and the total band 3 were both phosphorylated following incubation of intact cells with ^{32}P (Fig. 2). External cleavage of erythrocytes with α -chymotrypsin reduced both classes of band 3 to a characteristic 55,000-MW membrane-bound fragment^{23,34,35} which retained ^{32}P -label (Fig. 2). The ^{32}P -label was cleaved entirely from these fragments by a second digestion of inverted vesicles (not shown), indicating that it is localised on the cytoplasmic surface of the membrane. Thus, ankyrin-linked band 3 spans the membrane, and exhibits a significant similarity to the major band 3 protein.

A more detailed comparison of these polypeptides was provided by isoelectric focusing of cyanogen bromide fragments of SDS-PAGE-purified ankyrin-linked band 3 from immunoprecipitates and free band 3 from gel filtration (Fig. 4). The patterns are nearly identical for the two proteins, although relative peak heights differ (Fig. 4). This difference may be explained by the

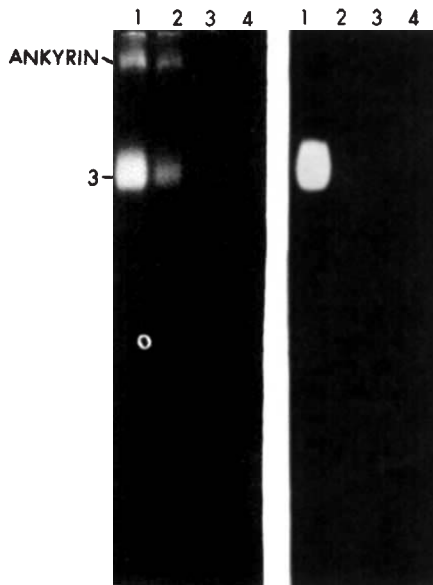


Fig. 5 Binding of purified spectrin to ankyrin-linked band 3 and to free band 3. 125 I-labelled samples from gel filtration of solubilised spectrin-depleted vesicles (Fig. 3) containing ankyrin and band 3–4.2 (left panel, lane 1) or band 3–4.2 free of ankyrin (right panel, lane 1) were incubated (90 min, 4 °C) with 75 $\mu\text{g ml}^{-1}$ purified unlabelled spectrin^{17,18} in the absence (lane 2) and presence (lane 3) of 100 $\mu\text{g ml}^{-1}$ unlabelled spectrin-binding fragment of ankyrin¹⁹ in column buffer (Fig. 3). Some of the samples (lane 4) had been acidified to pH 2 with HCl and neutralised before addition of spectrin. Spectrin– 125 I-protein complexes were precipitated by addition of monospecific anti-spectrin IgG (affinity purified) and protein A staphylococci, and the pellets were washed with column buffer. The immunoadsorbed proteins were analysed by SDS electrophoresis (Fig. 1) and visualised by autoradiography. The same amount of spectrin was precipitated in the presence or absence of the ankyrin fragment as determined from Coomassie blue staining (not shown).

fact that labelling was carried out with intact vesicles, and bound ankyrin could have altered the reactivity of portions of the band 3 polypeptide. Further evidence that ankyrin-linked band 3 is closely related to free band 3 is that antibodies raised against the 42,000-MW cytoplasmic domain of band 3 (ref. 35) cross-react with both polypeptides (in preparation).

Spectrin binds to ankyrin-linked band 3 but not to free band 3

Incubation of purified spectrin with 125 I-labelled ankyrin and band 3–4.2 isolated by gel filtration produced a spectrin–ankyrin–band 3–4.2 complex which was immunoprecipitated by monospecific anti-spectrin IgG (Fig. 5). The complex did not form if these proteins were exposed to acidic pH before addition of spectrin (Fig. 5), a treatment which destroys the spectrin-binding capacity of membranes¹⁸. The association of spectrin with these polypeptides was mediated by ankyrin, as the unlabelled spectrin-binding fragment of ankyrin reduced binding by 90% but had no effect on immunoadsorption of spectrin (Fig. 5). Furthermore, when isolated by gel filtration, band 3–4.2 (lacking ankyrin) did not bind to spectrin (Fig. 5). Other fractions from Fig. 3 also exhibited no binding to spectrin (not shown).

These results indicate that spectrin can reassociate with its membrane attachment site in the presence of Triton X-100. Consequently, it would be expected that the complex of spectrin–ankyrin–band 3–4.2 would be retained by the cytoskeletal assembly which remains after extraction of erythrocytes and ghosts with Triton X-100^{36,37}. Nearly all of spectrin and 10–15% of band 3 are recovered in Triton X-100-extracted ghosts along with bands 4.1, 4.2 and 7, and erythrocyte actin (Fig. 6). This residual band 3 is tightly adsorbed, and persists at a constant level during repeated washes in detergent (Fig. 6). Ankyrin is poorly visualised due to interference by spectrin, but is detected by radioimmunoassay in the same amounts as in whole ghosts

(not shown). Furthermore, ankyrin together with band 3, band 4.2 and band 7 can be extracted without solubilising these Triton residues³⁸ (Fig. 6). Cytoskeleton-linked band 3 is cleaved to a 55,000-MW fragment by external cleavage of erythrocytes with α -chymotrypsin before detergent extraction (Fig. 6). A second limited digestion of the cytoskeletons with α -chymotrypsin cleaves band 3 and the 55,000-MW fragment (not shown). It can thus be concluded that the cytoskeleton-bound band 3 spans the membrane and is the same polypeptide as the ankyrin-linked band 3.

Discussion

These data indicate that about 15% of band 3 is tightly linked to the erythrocyte cytoskeleton through association with ankyrin, with a stoichiometry of 1 mol band 3 per mol of ankyrin. The ankyrin–band 3 oligomer also contains a 72,000-MW polypeptide (band 4.2) which has been demonstrated to bind to band 3 in solution³⁰. Band 3 is present in about 10^6 copies per cell arranged in a dimeric form visualised as intramembrane particles³². Ankyrin, which is present in about 10^5 copies per cell, is thus associated with 10^5 band 3 molecules which are presumably also dimers and form from 5×10^4 to 10^5 intramembrane particles, depending on the ankyrin–particle stoichiometry. These ankyrin-linked structures are also the points of attachment of spectrin to the membrane, and may be nucleation sites for assembly of the cytoskeleton. The average nearest neighbour distance between ankyrin–band 3–4.2 oligomers would be 350–500 Å, which is comparable to the dimensions of tetrameric spectrin³⁹, the form of this protein in erythrocyte ghosts^{39,40}. Scanning electron microscopy of erythrocyte ghosts and cytoskeletons reveals a net-like appearance with globular structures 200–1,000 Å apart interconnected by filaments⁴¹. These dimensions are consistent with the possibility that the interstices of the network contain ankyrin-linked oligomers.

The finding that only a fraction of band-3 proteins are linked to spectrin helps reconcile two conflicting sets of data regarding association of integral proteins with the cytoskeleton in red cell membranes. Spectrin, and possibly actin, are thought to restrict lateral translation of intramembrane particles¹⁵ and to be associated with sialoglycoproteins¹⁶. Spectrin has been localised in

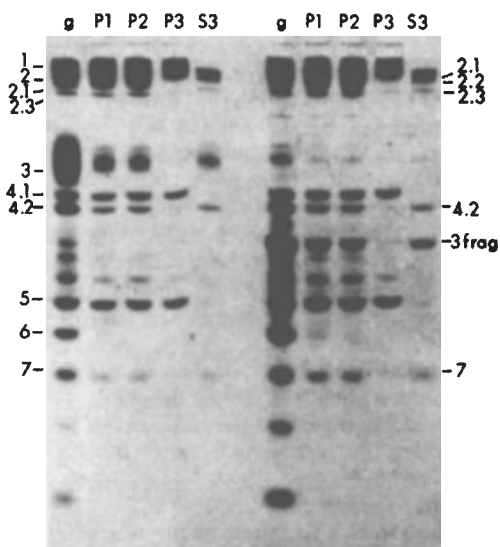


Fig. 6 Retention of band 3 in Triton X-100-extracted cytoskeletons of erythrocyte ghosts. Ghosts were prepared from erythrocytes that were untreated (left) or had been digested externally with α -chymotrypsin (right) as in Fig. 2. Cytoskeletons were prepared by incubating (15 min, 4 °C) the ghosts with 20 vol of 100 mM KCl, 10 mM NaPO₄, pH 7.5, 1 mM NaEDTA, 10 $\mu\text{g ml}^{-1}$ phenylmethylsulphonyl fluoride, 0.2 mM dithiothreitol and 0.2% (v/v) Triton X-100. Cytoskeletons were pelleted (15 min, 40,000g) (P1), washed once in extraction buffer (P2) and incubated (15 min, 0 °C) in extraction buffer plus 1 M KCl. The suspensions were pelleted, the supernatants (S3) removed and dialysed, and the pellets (P3) resuspended in the original volume. The samples were analysed by SDS electrophoresis (Fig. 1) and proteins stained with Coomassie blue.

patterns coincident with intramembrane particles⁴². The lateral mobility of integral membrane proteins is restricted, with a diffusion constant of $5 \times 10^{-12} \text{ cm}^2 \text{ s}^{-1}$ at 22 °C (refs 43, 44), and capping of lectins does not occur in erythrocytes⁴. These experiments have led to the view that intramembrane particles and their component proteins are attached to the cytoskeleton. However, no association could be demonstrated between spectrin and Triton X-100-solubilised band 3 or glycophorin³⁰. A direct role for band 3 and the sialoglycoproteins in the attachment of spectrin to the membrane has been ruled out by several experiments¹⁸⁻²⁰. The rotational diffusion of band 3 occurs at a rate which indicates that the membrane viscosity should support lateral diffusion at 10^{-9} – $10^{-10} \text{ cm}^2 \text{ s}^{-1}$ (ref. 45). Moreover, rotational diffusion of band 3 was unchanged following extraction of spectrin⁴⁵. Further evidence that integral proteins are not all immobilised is the formation of membrane vesicles depleted of spectrin and actin, but containing a full complement of integral proteins^{46,47}.

The data indicating that band 3 is unrestricted in the membrane are consistent with the present results that only ankyrin-associated band 3 is linked to spectrin. The fact that immobilisation of only a portion of band 3 proteins apparently leads to restriction of all these proteins suggests the involvement of an additional process. It is conceivable that the free proteins

are simply entangled in the underlying cytoskeletal meshwork. Another possibility is that the integral proteins associate with each other with low affinity. It is pertinent to this point that intramembrane particles are not dispersed randomly, but rather exist in groups of two to eight. The effect of such aggregation would be to decrease the diffusion constant due to increased size, and to immobilise particles associated with an ankyrin-linked particle.

These studies may have direct relevance to cells other than erythrocytes, as ankyrin has been detected by radioimmunoassay in many tissues and cell types⁴⁸. Future areas of interest include elucidation of the nature of the linkage between band 3 and ankyrin, the role of band 4.2 in this complex, and examination of possible association between ankyrin, band 3 and other membrane proteins such as hormone receptors.

We thank Dr Michael Sheetz for sharing data before publication which demonstrated that cytoskeleton-linked band 3 can be released by high salt extraction.

Note added in proof: Further evidence for association of a portion of the band 3 proteins with the cytoskeleton is provided by the finding that a small amount of band 3 can be chemically cross-linked to band 1 of spectrin in erythrocyte membranes⁵⁰.

Received 9 April; accepted 1 June 1979.

- Singer, S. J. & Nicolson, G. L. *Science* **175**, 720–731 (1972).
- Gitler, C. A. *Rev. Biophys. Bioengng* **1**, 51–86 (1972).
- Edelman, G. M. *Science* **192**, 218–226 (1976).
- Nicolson, G. L. *Biochim. biophys. Acta* **457**, 57–108 (1976).
- de Petris, S. & Raff, M. C. *Eur. J. Immun.* **2**, 523–535 (1972).
- Schlessinger, J., Schecter, Y., Willingham, M. C. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2659–2663 (1978).
- Anderson, R. G. W., Brown, M. S. & Goldstein, J. L. *Cell* **10**, 351–364 (1977).
- Korn, E. D. in *Cold Spring Harb. Conf. Cell Proliferation* Vol. 13 (eds Goldman, R., Pollard, T. & Rosenbaum, J.) 623–630 (Cold Spring Harbor Laboratory, New York, 1976).
- Rees, D. A., Lloyd, C. W. & Thom, D. *Nature* **267**, 124–128 (1977).
- Koch, G. & Smith, M. *Nature* **273**, 274–278 (1978).
- Flanagan, J. & Koch, G. *Nature* **273**, 278–281 (1978).
- Carter, W. R. & Hakomori, S. *J. biol. Chem.* **253**, 2867–2874 (1978).
- LaCelle, P. L. & Kirkpatrick, F. H. in *Erythrocyte Structure and Function* (ed. Brewer, G. J.) 535–557 (Liss, New York, 1975).
- Sheetz, M. P. & Singer, S. J. *J. Cell Biol.* **73**, 638–646 (1977).
- Elgsaeter, A. & Branton, D. *J. Cell Biol.* **63**, 1018–1036 (1974).
- Nicolson, G. L. & Painter, R. G. *J. Cell Biol.* **59**, 395–406 (1973).
- Bennett, V. *Life Sci.* **21**, 433–440 (1977).
- Bennett, V. & Branton, D. *J. biol. Chem.* **252**, 2753–2763 (1977).
- Bennett, V. *J. biol. Chem.* **253**, 2292–2299 (1978).
- Bennett, V. & Stenbuck, P. J. *J. biol. Chem.* **254**, 2533–2541 (1979).
- Luna, E., Kidd, G. K. & Branton, D. *J. biol. Chem.* **254**, 2526–2532 (1979).
- Fowler, V. & Bennett, V. *J. supramolec. Struct.* **8**, 215–221 (1978).
- Bretscher, M. S. *J. molec. Biol.* **59**, 351–357 (1971).
- Boxer, D. H., Jenkins, R. E. & Tanner, M. J. A. *Biochem. J.* **137**, 531–534 (1974).
- Steck, T. L. *J. Cell Biol.* **62**, 1–19 (1974).
- Kessler, S. J. *Immunology* **115**, 1617–1624 (1975).
- Yu, J. & Goodman, S. *Fedn Proc.* **38**, 798a (1979).
- Shapiro, D. L. & Marchesi, V. T. *J. biol. Chem.* **252**, 508–519 (1977).
- Blumenfeld, O. O., Gallop, P. M. & Liao, T. H. *Biochem. biophys. Res. Commun.* **48**, 242–251 (1972).
- Yu, J. & Steck, T. L. *J. biol. Chem.* **250**, 9176–9184 (1975).
- Steck, T. L. *J. molec. Biol.* **66**, 295–305 (1972).
- Yu, J. & Branton, D. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3891–3895 (1976).
- Nigg, E. & Cherry, R. J. *Nature* **277**, 493–494 (1979).
- Jenkins, R. E. & Tanner, M. J. A. *Biochem. J.* **147**, 393–399 (1975).
- Steck, T. L., Ramos, B. & Strapazon, E. *Biochemistry* **15**, 1154–1161 (1976).
- Yu, J., Fischman, D. A. & Steck, T. L. *J. supramolec. Struct.* **1**, 233–248 (1973).
- Sheetz, M. P. & Sawyer, D. J. *supramolec. Struct.* **8**, 399–412 (1978).
- Sheetz, M. P. *Biochim. biophys. Acta* (in the press).
- Ralson, G. B. *J. supramolec. Struct.* **8**, 361–373 (1978).
- Ungewickell, E. & Gratzner, W. *Eur. J. Biochem.* **88**, 379–386 (1978).
- Hainfeld, J. F. & Steck, T. L. *J. supramolec. Struct.* **6**, 301–311 (1977).
- Shotton, D., Thompson, K., Wofsy, L. & Branton, D. *J. Cell Biol.* **76**, 512–531 (1978).
- Peters, R., Peters, J., Tews, K. H. & Bähr, W. *Biochim. biophys. Acta* **367**, 282–294 (1974).
- Fowler, V. & Branton, D. *Nature* **268**, 23–26 (1977).
- Cherry, R. J., Bürkli, A., Busslinger, M., Schneider, G. & Parish, G. R. *Nature* **263**, 389–393 (1976).
- Lutz, H. U., Liu, S.-C. & Palek, J. *J. Cell Biol.* **73**, 548–560 (1977).
- Allan, D., Billah, M. M., Finean, J. B. & Michell, R. H. *Nature* **261**, 58–60 (1976).
- Bennett, V. (in preparation).
- Fairbanks, G., Steck, T. L. & Wallach, D. F. H. *Biochemistry* **10**, 2606–2617 (1971).
- Liu, S.-C. & Palek, J. *J. supramolec. Struct.* **10**, 97–109 (1979).

letters

A precessing ring model for SS433

THE wavelengths of the shifted optical lines in SS433 seem to show a periodic variability at a period of 160 ± 3 d (or a multiple integer thereof)^{1,2}. We have suggested that such periodicity could arise if the lines were emitted from a precessing ring of matter orbiting around a massive black hole (BH) of mass M (ref. 3). We also suggested that SS433 may be located inside an old globular cluster³. We describe here, in principle, how such precession model may come about and what some of its observational constraints might be.

We suggest that a normal, solar type, star has been captured⁴ by a $\sim 10^6 M_\odot$ BH. Its orbit then circularises at the tidal radius⁵, which is located at $R \sim 45 R_G$ ($R_G \equiv GM/c^2$), as is indicated by the wavelengths of the shifted lines^{3,2}. Note that the tidal radius for this captured star, of average density ρ , is given by⁶

$$R_T \sim 45 \left(\frac{\rho}{1 \text{ g cm}^{-3}} \right)^{-1/3} \left(\frac{M}{3 \times 10^6 M_\odot} \right)^{-2/3} R_G$$

and that the lifetime of that orbit against gravitational radiation is⁵

$$(R/\dot{R})_T \equiv \tau \sim 10^6 \left(\frac{M}{3 \times 10^6 M_\odot} \right)^{-2/3} \text{ yr}$$

(The lifetime may be shorter when appreciable gas dissipation occurs).

We now assume that this star experiences steady Roche-lobe overflow due to either self expansion or steady drift inwards. Once overflow commences, the particle jet through L1 follows an orbit which is almost identical to that of the star, because of the relative smallness of the stellar Roche-lobe radius (this is unlike some 'normal' X-ray or dwarf novae binaries, where the jet goes right into the other lobe). The jet thus forms a cool circular keplerian ring (or torus) around the BH, of small width

$$\sim 0.04 R \left(\frac{r}{10^{11} \text{ cm}} \right) \left(\frac{3 \times 10^6 M_\odot}{M} \right)$$

where r is the stellar radius. This ring is steadily controlled by the L1 overflow.

We assume the BH to be rotating, not necessarily at extreme Kerr frequencies, because such massive BHs tend to reduce their a/M when 'swallowing' stars⁷. If the star is orbiting in an inclined orbit, a precession of its orbital plane around the rotation axis will occur due to the spin-orbit coupling, at a period given by⁸

$$T_p \sim 48.6 \text{ d} \left(\frac{a}{M} \right)^{-1} \left(\frac{M}{3 \times 10^6 M_\odot} \right) \left(\frac{R}{45 R_G} \right)^3$$

This period will be consistent with the value of 160 ± 3 d for $a/M \sim 1/3(M/3 \times 10^6 M_\odot)$. (Note that a mass $M > 10^7 M_\odot$ will not be consistent with this model if the period is, indeed, 160 d).

A similar precession motion is also performed by the ring particles, at slightly different frequencies; however, as the ring is always additionally controlled by the L1 overflow, with a time scale of ~ 100 orbital revolutions (≥ 10 d), it moves essentially with the stellar precession, having a 'thickness' comparable to its width. This time scale is, essentially, the time it takes for a matter blob leaving L1 to separate from L1 by about an orbital length, and the 'thickness' results from the precession frequency gradient across the ring.

Frictional processes in the ring will bring about a slow drift inwards of 'old' ring material and, eventually, because of dephasing of its precessional motion, its settling down to the outer ring of an equatorial disk⁹, which is where all of these out-of-phase inclined inner rings intersect. The disk may have pre-existed around the BH³. It provides continuum UV and optical flux to excite the 'new', cold, ring (jet) material into line emission. As the 'old' material drifts inwards, its density first decreases in the diffusion process and then it heats up on approaching the disk. Both these effects should secure the uniqueness of that ring, relative to all material inside it, as an optical line emitter. Note that since the inside material tends more and more to lie in the BH equator⁹, the ring is likely to 'see' the inside UV source without much interruption from intervening opacity. The volume of the ring is $\sim 10^{37} \text{ cm}^3$, requiring a density of $\geq 5 \times 10^{11} \text{ cm}^{-3}$ (ref. 10) to account for the observed line luminosity. A cool ring of such dimensions and density and with solar type composition will indeed be optically thick to UV radiation but marginally optically thin in the visible when viewed pole-on.

To analyse observational constraints on our model, we write the centre of mass (COM) of the lines³ as

$$\frac{1}{2} \frac{v^2}{c^2} + \frac{\phi}{c^2} = \frac{3}{2\eta^2} \frac{v^2}{c^2} \approx 0.033 \quad (\text{ref. 2})$$

with ϕ the gravitational potential and v the velocity of the emitting region. For ring geometry $\eta^2 = 1$, while $\eta^2 \sim 3$ for $v^2 \gg \phi$.

Recent data of Abell and Margon¹¹ suggest that the ring radiation geometry should be such that the ring does not radiate homogeneously but rather radiates predominantly from the regions which lie close to the BH equator, possibly due to the hot outer ring of the disk¹⁵.

This geometry is equivalent to the double-water-hose kinematic geometry of Milgrom² and Abell and Margon¹¹, except that for the ring, $\eta^2 = 1$ while Abell and Margon take $\eta^2 = 3$, (that is, they use $v/c = 0.27$ while we have $v/c \approx 0.15$). On repeating the Abell and Margon¹¹ analysis for $\eta^2 = 1$, we find that

$$(i, \theta) \equiv (66^\circ, 32^\circ) \text{ or } (32^\circ, 66^\circ)$$

where θ is the angle between the plane of the ring and the BH rotation axis, and i is the angle between the line of sight and the BH rotation axis.

The width of the shifted lines emitted from these nodal regions depends on the length of these regions and on the thickness of the ring. Note that since the intrinsic intensity distribution per unit Doppler-shifted frequency along such ring segment maximises, almost always, at one end³, and since the

illumination should maximise at about the middle of the segment, one may sometimes expect double-peaked lines to be observed, depending on the relative heights of these maxima. Furthermore, as the optical depth through which the segment is viewed in the lines increases when the line of sight becomes more parallel to it, one also expects intensity variations with precessional phase.

Finally, we mention briefly a possible origin for the unshifted lines in the spectrum of SS433, which show a variability over time scale of a day¹², self-absorption and widths of $\approx 1,000 \text{ km s}^{-1}$. Such variability limits the emission region to $\leq 2 \times 10^{15-16} \text{ cm}$. For a BH of $\sim 10^6 M_\odot$, velocities of a few thousand km s^{-1} are compatible with such distances. If SS433 is accreting clouds, possible in the SNR W50 (ref. 3), moving at $\geq 100 \text{ km s}^{-1}$, or else is accreting interstellar gas³ moving at $\geq 10 \text{ km s}^{-1}$,

$$\left[\left(\frac{n}{1 \text{ cm}^{-3}} \right) \left(\frac{v}{10 \text{ km s}^{-1}} \right)^{-3} \sim 0.01 - 1 \right],$$

then cooling in the accreted matter will be efficient down to $\sim 10^{15-16} \text{ cm}^3$. In that region, density and temperature may be adequate to emit most of the unshifted emission lines, and its illumination and velocity distribution, modified by angular momentum of infalling matter¹³, may cause the line width and shape (see for example ref. 14). Note that the interior part of the accreting matter may become sufficiently hot to radiate at X-ray wavelength, and that its influence on the kinematics and structure of the ring will be small due to the high, supersonic, velocities. It may, however, deposit by electrons up to $\sim 3 \times 10^{36} \text{ erg s}^{-1}$ in the ring, to add to its excitation.

We thank A. Dekel, J. Katz and Y. Tuchman for discussions. T.P. thanks the NSF for support (through grant PHY77-22489).

A. AMITAI-MILCHGRUB*

T. PIRAN†

J. SHAHAM*

* *Racah Institute of Physics,
The Hebrew University of Jerusalem,
Jerusalem, Israel*

† *Center for Relativity,
The University of Texas,
Austin, Texas 78712*

Received 14 May; accepted 4 July 1979.

1. Margon, B. *et al.* *IAU Circ.* No. 3345 (1979).
2. Milgrom, M. *Astr. Astrophys. Lett.* (in the press).
3. Amitai-Milchgrub, A., Piran, T. & Shaham, J. *Nature* **279**, 505 (1979).
4. Fabian, A. C., Pringle, J. E. & Rees, M. J. *Mon. Not. R. astr. Soc.* **172**, 15P-18P (1975).
5. Frank, J. & Rees, M. J. *Mon. Not. R. astr. Soc.* **176**, 633-647 (1976).
6. Hills, J. G. *Nature* **254**, 275-298 (1975).
7. Young, P. J. *Astrophys. J.* **212**, 227-233 (1977).
8. Wilkins, O. C. *Phys. Rev. D* **5**, 814-822 (1972).
9. Bardeen, J. M. & Petterson, J. A. *Astrophys. J. Lett.* **195**, L65-L67 (1975).
10. Osterbrock, D. E. *Astrophysics of Gaseous Nebulae*, 59-96 (Freeman, San Francisco, 1974).
11. Abell, G. O. & Margon, B. *Nature* **279**, 701-703 (1979).
12. Margon, B. *et al.* *Astrophys. J. Lett.* **230**, L41-L45 (1979).
13. Shvartsman, V. F. *Sov. Astr.* **15**, 377-384 (1971).
14. Ulrich, R. K. *Astrophys. J.* **210**, 377-391 (1976).
15. Amitai-Milchgrub, A. & Shaham, J. *Astr. Astrophys. Lett.* (in the press).

On the nature of SS433

THE report of large wavelength variations associated with SS433¹ illustrates the large amount of information to be gleaned from preliminary data when effects large enough to involve the theory of relativity are present. It seems most plausible to interpret the large periodic wavelength variations in the Balmer lines of this object as arising from periodic motion, requiring, as a result of its magnitude, the special relativistic Doppler formula for its interpretation. Here we comment on some recent observations that have made it possible to make quantitative estimates of some aspects of the model.

If the motion may be characterised by a constant speed, then the classical Doppler effect contains information about the radial velocity, while the special relativistic form also involves

the actual speed. The relativistic expressions for these parameters are

$$\frac{v \sin i}{c} = (z_2 - z_1)/(2 + z_1 + z_2) \quad (1)$$

and
$$\frac{v}{c} = [1 - 4/(2 + z_1 + z_2)^2]^{1/2}$$

where z_1 and z_2 are the minimum and maximum values respectively for the observed redshift relative to the centre of mass. Using the values given for the object in ref. 1, one finds that it may be characterised by a velocity of $80,000 \text{ km s}^{-1}$, inclined 30° to the plane of the sky.

Since the object exhibits both X rays² and spectral lines of hydrogen and neutral helium¹, it seems reasonable to suggest that the latter are formed far from the ionising effects of the former. Thus any attempt to achieve the observed velocities by dynamical means incorporating the observed period of 160 d¹ requires a mass in excess of 10^8 solar masses. If the distance of 3.5 kpc (ref. 3) is correct then the presence of such a massive object is incompatible with stellar dynamics in that region of the Galaxy. Thus it is plausible to suggest that some form of hydromagnetic acceleration is responsible for the extremely large velocity. It has also almost become commonplace to suggest accretion by a collapsed object as the source of X rays in other systems exhibiting periodic variations.

These two suggestions strongly imply that a stellar binary system exists, one component being a rather massive star possessing a strong magnetic field (not unlike VV Cephei⁴), and the other a collapsed object, which is indeed the source of the X-ray spectrum. The rapid mass loss from the star would force matter not only through the inner lagrangian point but also through the outer lagrangian points, into the magnetic field. Except for the immediate X-ray shadow cast by the star, this material would be fully ionised by the X-ray source, and thus be frozen to the co-revolving magnetic field of the star. This ionised material would be accelerated outwards until it exceeds the Alfvén velocity or passes the Strömgren radius dictated by the X-ray source. The outward motion could be expected by means of cyclotron radiation to produce the observed non-thermal radio radiation⁵. At the Strömgren radius, the material will radiate and rapidly cool, giving rise to the observed high-velocity cool gas spectrum.

Additional observational data have recently become available which enable us to make quantitative estimates of several aspects of this model. The observed H α flux of Liebert *et al.*⁶, the 3.5 kpc distance, and the reddening of Margon *et al.*³ imply an emitted energy at H α of about $3 \times 10^{34} \text{ erg s}^{-1}$ or 5×10^{45} H α photons per s per stream. If β is the number of H α photons emitted for each outflowing proton, then the implied mass outflow would be 10^{-4} per β solar masses per yr. The divergence of the field would allow the material streaming outwards to spread, so that radiation is produced in a large volume. If the emitting region subtends about 15° as seen from the binary, then the time required for this region to cross the location of maximum Doppler shift would be about 4% of the 160 d period, and would produce the flat tops of about one week duration of the radial velocity curves¹. The rotationally outflowing material will compress the field in the rotational plane over that of a static dipole geometry, in a manner perhaps similar to that of the Sun or Jupiter. As the distance from the binary system increases, the weakening field will form a spiral, which when combined with the above compression, yields a field decreasing in strength much more slowly than a static dipole field. This phenomenon is well observed in both the Sun and Jupiter, and implies a field strength that varies with distance as $r^{-\alpha}$, with $\alpha \approx 1.3$ (ref. 7). The spiraling of the magnetic field lines implies that the mass motion at time of radiation will have both tangential and radial components. If we take the magnetic and kinetic energies to be equal in the radiating volume, and equating the radial and tangential velocity components, then assuming co-rotation, the distance to the emitting regions is fixed. In addition, the previously calculated mass flux will specify the density, and hence

the magnetic field strength, in the emitting region. If we extrapolate the field to the stellar surface using the above power law dependence, we may place a lower limit on the mass of the star such that it is stable against magnetic distortion⁴. The expression that relates these parameters is

$$\frac{M_*}{M_\odot} \geq \frac{10^{-8} v^{\alpha+1} P^\alpha \sin^\alpha \xi \left(\frac{R_*}{R_\odot}\right)^{2-\alpha} \rho^{1/2}}{R_\odot (2\pi)^{\alpha-1/2}} \quad (2)$$

where ξ is the radial angle of the recombining gas, which we have taken to be 45° . With the previously determined values of α , period P and velocity v , this becomes

$$\frac{M_*}{M_\odot} \geq 1.4 \times 10^9 (R_*/R_\odot)^{0.7} \rho^{1/2} \text{ cgs} \quad (3)$$

The geometrical arguments previously mentioned and the outflowing particle flux yield a density in the beam of

$$\rho = 4\pi N m_H (\beta v^3 P^2 \theta^2 \cos \xi \sin^2 \xi)^{-1} \approx 1.6 \times 10^{-19} / \beta \text{ g cm}^{-3} \quad (4)$$

where N is the number of H α photons emitted per s and θ is the half-angle subtended by the emitting regions as seen from the binary system. We may take the keplerian separation implied by the large mass indicated in equation (3), and the period P , to suggest an upper limit to the stellar radius of 2 AU. Equation (3) implies a lower limit to the mass of the primary of $40 M_\odot / \beta^{1/2}$.

The absence of forbidden lines in the spectrum of SS433 and the recombination time for the hydrogen spectrum require a high-density medium for the emitting region. In order to keep the kinetic energy loss through the beam within reasonable bounds, one would expect the beam density to be low. These two requirements suggest a high speed particle beam interacting with a high density medium at the boundary of the Stromgren sphere. Although the details of such an interaction are complicated, mechanisms such as charge exchange imply that a value of β very much larger than that obtained from classical recombination theory would result. A large value of β implies that typical stellar masses would be gravitationally stable against the required stellar magnetic field. The evidence that the object contains a massive blue star⁸ indicates that the field, although large, can be well below the disruption strength. Also a large value of β reduces the required mass loss rate to values typical for massive stars.

Thus we see that many aspects of this object are consistent with a model containing a massive blue star of the order of 50 solar masses in orbit with a collapsed object, separated by about 3 AU. The strong magnetic field of this blue star declines slowly with distance to a value of about 1 G at 1,000 AU, where the outflowing matter interacts with the surrounding medium. Note that this distance, which is largely set by the geometry of the model, is consistent with the Strömgren radius expected for the X-ray source² and densities we have used.

Recently Abell and Margon⁹ have shown that the shape of the observed segment of the radial velocity curve implies that the expanding streams may be inclined to the orbital plane at the location of emission. If confirmed, this interpretation would imply that the stellar field be inclined to the orbital plane, introducing only minor modifications to our model.

G. W. COLLINS II
G. H. NEWSOM

Department of Astronomy,
The Ohio State University,
Columbus, Ohio 43210

Received 30 April; accepted 19 June 1979.

1. Margon, B. *et al.* *IAU Circ.* No. 3345 (1979).
2. Marshall, F. E., Mushotzky, R. F., Boldt, E. A., Holt, S. S. & Serlemitsos, P. J. *IAU Circ.* No. 3314 (1978).
3. Margon, B., Ford, H. C., Katz, J. I., Kwitter, K. B. & Ulrich, R. K. *Astrophys. J.* **230**, L41 (1979).
4. Chandrasekhar, S. & Fermi, E. *Astrophys. J.* **118**, 116 (1953).
5. Seaquist, E. R., Gregory, P. C. & Crane, P. C. *IAU Circ.* No. 3256 (1978).
6. Liebert, J., Angel, J. R. P., Hege, E. K., Martin, P. G. & Blair, W. P. *Nature* **279**, 384 (1979).
7. Brandt, J. C. *Introduction to the Solar Wind*, 132 (Freeman, San Francisco, 1970).
8. Clark, D. H. & Mordin, P. *Nature* **276**, 44 (1978).
9. Abell, G. O. & Margon, B. Preprint (1979).

Cool gaseous nebulae

THE electron temperatures of diffuse gaseous nebulae have long been thought to be close to 10^4 K. Much lower temperatures were derived from some of the early radio continuum and recombination line work, but these were generally considered to be wrong for a variety of reasons (cf. the review by Seaton¹). However, most of these considerations applied only to the easily measurable bright nebulae like Orion, and these are expected to be at high temperatures because of their high densities and the effects of collisional de-excitation on cooling agents such as oxygen and neon ions². While there is little doubt that the bright nebulae do indeed have temperatures of $\sim 8,000$ – $9,000$ K (refs 3–5), there are strong indications that some nebulae of lower densities have much lower temperatures, $\leq 4,000$ – $5,000$ K (ref. 5). We have made radio recombination line observations of several low-density nebulae, in order to determine electron temperatures in the absence of such effects as collisional de-excitation, stimulated emission, and pressure broadening. Several of these nebulae have been found to have temperatures below 5,000 K, and for two of them which are discussed here (RCW94 and G339.1–0.2) absolute upper limits of $\sim 4,700$ K are imposed by the line widths alone.

The observations were made at 5 GHz with the Parkes 64-m radio telescope in March and April 1979. The system temperature was about 50 K, and the half-power beamwidth 4.4 ± 0.1 arc min. The temperature scale was based on observations of Hydra A, assumed to have a flux density of 13.1 Jy; the ratio of point source flux density to main beam brightness temperature was 1.43. The digital correlation spectrometer provided two simultaneous 512-channel spectra of a 10-MHz band containing the H 109 α , H 137 β , and He 109 α lines; the spectral resolution was 2.3 km s^{-1} after Hanning smoothing. The total power observing mode was used; after each on-source integration,

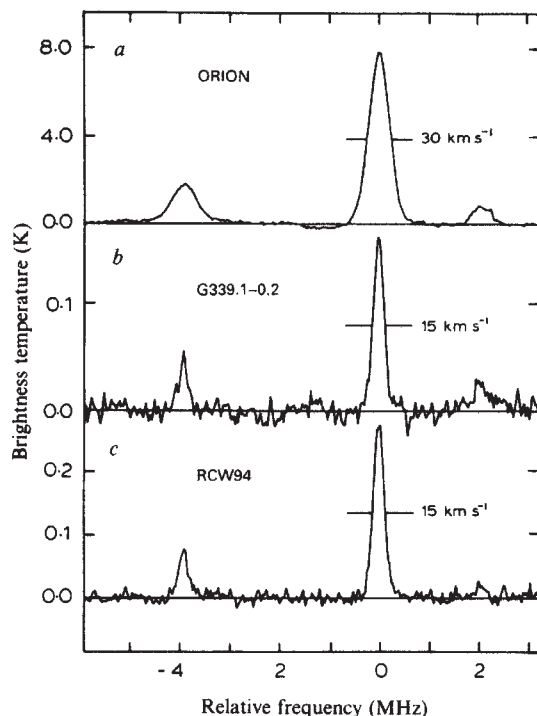


Fig. 1 5-GHz recombination line spectra of the Orion Nebula (a), G339.1–0.2 (b) and RCW94 (c). Hanning smoothing has been applied. A quadratic baseline was removed from the spectrum of the Orion Nebula and linear baselines from those of G339.1–0.2 and RCW94. The frequency scale is relative to the H 109 α line; the H 137 β and He 109 α lines appear at -3.890 and $+2.041$ MHz respectively.

Table 1 Observed and derived parameters

	RCW94	G339.1–0.2
α (1950)	15 h 38 min 13 s	16 h 40 min 48 s
δ (1950)	$-54^{\circ}06.7'$	$-45^{\circ}58.5'$
T_c (K)	1.69 ± 0.09	0.85 ± 0.05
H 109 α :		
ΔT_L	0.273 ± 0.005	0.160 ± 0.005
ΔV_L (km s^{-1})	14.9 ± 0.2	14.9 ± 0.3
V_L (km s^{-1})	-43.1 ± 0.2	-119.7 ± 0.2
$I(\text{H } 137\beta)/I(\text{H } 109\alpha)$	0.27 ± 0.02	0.28 ± 0.02
$I(\text{He } 109\alpha)/I(\text{H } 109\alpha)$	0.06 ± 0.02	0.10 ± 0.02
T_D (K)	$4,730 \pm 130$	$4,730 \pm 190$
T_e^* (K)	$4,660 \pm 280$	$3,960 \pm 280$
T_e (K)	$4,490 \pm 280$	$3,800 \pm 280$

Quoted errors are 1 s.d. and allow for baseline uncertainties. Continuum (T_c) and peak line (ΔT_L) temperatures are in units of brightness temperature averaged over the main beam. The linewidth (ΔV_L) is the FWHM value, and the line velocity (V_L) is with respect to the local standard of rest. The integrated line intensities (I) were obtained by integrating under the observed line profiles. The Doppler temperature (T_D) allows for the instrumental resolution of 2.3 km s^{-1} . T_e^* and T_e are electron temperatures computed from the line-to-continuum ratio; LTE was assumed in determining T_e^* , and T_e takes into account all non-LTE effects assuming a local electron density of 100 cm^{-3} .

lasting 20 min, a reference integration of the same duration was made over the same range of hour angle. Special care was taken to assure compatibility of the continuum and line observations; the continuum measurements were interspersed with the line measurements of a given source, and were made with long scans through the peak of the source in order to separate the source from the galactic background accurately. The final spectra of the two sources RCW94 and G339.1–0.2 are shown in Fig. 1, compared with a spectrum of the Orion Nebula; measured and derived parameters are given in Table 1.

The most striking features of these lines are their widths, both 14.9 km s^{-1} . This is extremely small compared with the widths of recombination lines from other gaseous nebulae; the average of 128 lines observed in one survey⁶ was 34.5 km s^{-1} , and the Orion Nebula with a linewidth of 30 km s^{-1} is typical of bright nebulae. These narrow lines almost certainly originate within the nebulae themselves, and not from cold partially ionised regions which may surround the nebulae; supporting evidence includes the absence of any significant carbon recombination line emission, approximately normal helium-to-hydrogen line intensity ratios, and β/α ratios close to the local thermodynamic equilibrium (LTE) value of 0.276. In addition the observed H 109 α profiles are well represented by single gaussians; if some fraction of the line emission were supposed to originate outside the nebulae, the central velocities of the two line components would have to agree within $\sim 0.5 \text{ km s}^{-1}$, and the widths and intensities would have to be interrelated to an improbable degree to satisfy the tight condition that the overall line shape is nearly gaussian. The linewidths thus imply an absolute upper limit of $4,730 \text{ K}$ (T_D in Table 1) for the electron temperatures of the nebulae. If there are any turbulent mass motions, the electron temperatures must be smaller still.

The electron temperatures determined from the line-to-continuum ratios are consistent with these upper limits. Assuming LTE, they are $4,660$ and $3,960 \text{ K}$ for RCW94 and G339.1–0.2 respectively. Corrections for non-LTE effects reduce these temperatures still further, to $4,490$ and $3,800 \text{ K}$ respectively assuming a local electron density of 100 cm^{-3} (the corresponding β/α ratio is 0.29 in both cases); these corrections include stimulated emission due to the continuum from the nebulae, the galactic background, and the 2.7 K microwave background (about 5% of the total line emission). The r.m.s. electron density is ~ 20 – 30 cm^{-3} for both nebulae, so the adopted local value corresponds to filling factors of ~ 0.05 – 0.10 , typical values; use of the r.m.s. densities gives the lowest possible values for the electron temperatures, $4,130$ and $3,530 \text{ K}$ for RCW94 and

G339.1–0.2 respectively, but the temperatures obtained using a density of 100 cm^{-3} are probably more realistic. The corresponding r.m.s. turbulent velocities are small, $\sim 3\text{--}5\text{ km s}^{-1}$, as might be expected for well evolved, low-density nebulae such as these.

The low electron temperatures probably require relatively high metal abundances^{7,8}. The normal helium-to-hydrogen line intensity ratios suggest that the nebulae are fully ionised. The excitation parameters (90 and 60 pc cm⁻² for RCW94 and G339.1–0.2 respectively) correspond to early O-stars; the exciting star of RCW94 has been identified as LS3386 (ref. 9), and recent near-IR measurements of it (to be published) are consistent with an O5 star at the distance of the nebula (2.7 kpc), with $A_V \sim 5$. And recent work indicates that dust which may be present in the nebulae could not produce the necessary cooling¹⁰. The most likely explanation for the low temperatures is therefore a high metal abundance; indeed, recent radio and optical observations of another low-temperature nebula, G339.2–0.4, reveal metal abundances higher than the solar values by a factor of 4–5 (ref. 11).

We thank Lynette M. Newton and Dominic M. Zarro for assistance with the observations and data reduction. P.A.S. thanks the staff of the CSIRO Division of Radiophysics for their hospitality.

Note added in proof: New optical observations of RCW94 (in preparation) reveal metal abundances higher than the solar values by a factor of 6–7.

P. A. SHAVER

European Southern Observatory, c/o CERN,
CH-1211 Geneva 23, Switzerland

R. X. MCGEE

CSIRO Division of Radiophysics,
PO Box 76, Epping, NSW 2121, Australia

S. R. POTTASCH

Kapteyn Astronomical Institute,
Postbus 800, 9700 AV Groningen, The Netherlands

Received 4 June; accepted 2 July 1979.

1. Seaton, M. J. *Q. J. R. astr. Soc.* **15**, 370 (1974).
2. Osterbrock, D. E. *Astrophys. J.* **142**, 1423 (1965).
3. Peimbert, M. & Torres-Peimbert, S. *Mon. Not. R. astr. Soc.* **179**, 217 (1977).
4. Wilson, T. L., Biegging, J. & Wilson, W. E. *Astr. Astrophys.* **71**, 205 (1979).
5. Shaver, P. A. *Astrophys. Lett.* **5**, 167 (1970).
6. Wilson, T. L., Mezger, P. G., Gardner, F. F. & Milne, D. K. *Astr. Astrophys.* **6**, 364 (1970).
7. Burbidge, G. R., Gould, R. J. & Pottasch, S. R. *Astrophys. J.* **138**, 945 (1963).
8. Balick, B. & Sneden, C. *Astrophys. J.* **208**, 336 (1976).
9. Johnson, H. M. *Publ. astr. Soc. Pacific* **85**, 586 (1973).
10. Sarazin, C. L. *Astrophys. J.* **211**, 772 (1977).
11. Shaver, P. A., McGee, R. X., Murdin, P. G. & Goss, W. M. *Astr. Astrophys.* (submitted).

Table 1 Parallax determinations with r.m.s. error $<10^{-2}$ arc s

Pulsar	Parallax (p) (10^{-3} arc s)	Dispersion measure (D) (pc cm ⁻³)	Electron density (n_e) (cm ⁻³)
0329+54	1.3 ± 1.7	26.8	0.035 ± 0.046
0809+74	1.8 ± 7.1	5.8	0.010 ± 0.041
0950+08	8.3 ± 8.8	3.0	0.025 ± 0.026
1237+25	-6.2 ± 5.2	9.3	0.058 ± 0.048
1818-04	3.1 ± 6.1	84.4	0.262 ± 0.515
1929+10	21.5 ± 8.0	3.2	0.069 ± 0.026
2016+28	0.9 ± 3.6	14.2	0.013 ± 0.051
2020+28	-1.3 ± 4.6	24.6	-0.032 ± 0.113
2021+51	1.8 ± 4.9	22.6	0.041 ± 0.111

and a 25-m telescope at Defford, a baseline of 127 km length. The technique involved a search for apparent motion between the pulsar and one or two nearby small diameter radio sources, assumed to be extragalactic and much more distant than the pulsar. Observations were made over a period of between 2 and 4 yr, each pulsar being observed once or twice a year.

A least squares fitting process was used to determine linear proper motions in right ascension and declination and the parallax. For many of the pulsars most of the observations were made when the Earth was in approximately the same part of its orbit around the Sun so that the parallactic motions were small and consequently the errors in the measured parallaxes large. However, for nine pulsars out of the 26 studied, data were collected sufficiently often to allow determinations of parallax with r.m.s. errors of <0.01 arc s and these are listed in Table 1. The measured proper motions are not included and the complete results for all 26 sources will be published elsewhere. The major errors in the parallax determinations, apart from the random errors caused by poor signal-to-noise ratios and limited sampling through the year, are due to secular variations in the properties of the ionosphere. Systematic errors are to be expected because of the solar cycle and annual ionospheric variations above the observatories. There are not sufficient ionospheric data to make accurate corrections to the observations but a computer analysis using plausible data indicates that these errors amount at most to a few milliarcseconds per year in proper motion and a few milliarcseconds in parallax. This is confirmed by the small parallaxes measured on some of the pulsars which have large dispersion measures and are therefore expected to be very distant. The dispersion measures are shown in Table 1. The errors in these values are all $<1\%$ (ref. 4).

Taken individually, only one pulsar, PSR1929+10, shows a significant parallax of 21.5 ± 8.0 mas. This pulsar has the second smallest dispersion measure of all known pulsars and the parallax places the pulsar at a distance of about 50 pc from the Sun. The mean electron density (n_e) along the line of sight to the pulsar can be obtained from the product of the parallax (p) and the dispersion measure (D). That is

$$n_e = p.D$$

$$\text{In this case, } n_e = 0.069 \pm 0.026\text{ cm}^{-3}$$

This is somewhat larger than the mean electron densities measured along the lines of sight to pulsars over large distances through the Galaxy determined from observations of H I absorption. However, this pulsar lies at a low galactic latitude ($b = -4^\circ$) and the line of sight to the pulsar is always within the disk of the Galaxy where the mean electron density is probably enhanced somewhat by material and ionisation associated with the population I stars and H I and H II regions. For pulsars, the galactic height distribution is several times wider than this enhancement, so that most of the lines of sight lie within this layer for only a small fraction of their lengths.

The errors in the parallax determinations for the other pulsars in Table 1 are too large to allow meaningful distances to be estimated. However, because the derived value of n_e is proportional to parallax, it is possible to obtain an unbiased estimate of the mean electron density for each pulsar and this is shown in

Measurements of the trigonometric parallax of pulsars

THE distances of pulsars from the Earth have been determined using the dispersion measure and an assumed constant value of the mean electron density through the Galaxy of about 0.03 cm^{-3} . Independent distance estimates are available for a number of distant pulsars from the measurement of the H I absorption in the pulsar radiation and these results are in general accord with the dispersion measure estimates over large distances through the Galaxy. There is some indirect evidence^{1,2} that the mean electron density in the solar neighbourhood is typical of the rest of the Galaxy but all estimates of pulsar distances within 1 kpc of the Sun rely on this assumption. We report here measurements of trigonometric parallax of pulsars and from the distance estimates we obtain a value for the mean electron density within about 500 pc of the Sun.

The observations formed part of the programme at Jodrell Bank for the measurement of the proper motions of pulsars³. Observations were made at 408 MHz using a radio-link interferometer between the 76-m MkIA telescope at Jodrell Bank

Table 1. Combining these results we can obtain a weighted mean of $0.029 \pm 0.014 \text{ cm}^{-3}$.

There is no strong evidence to suggest any inhomogeneity in the electron distribution, as the individual estimates of the electron density are all substantially in accord with the mean value of about 0.03 cm^{-3} when their errors are taken into consideration.

This value for the mean electron density is derived essentially from those pulsars within about 500 pc of the Sun (that is, dispersion measures $<15 \text{ pc cm}^{-3}$) since the estimates of n_e for the more distant objects all have large errors and therefore carry little weight in the formation of the mean. This value of the mean electron density is consistent with estimates over much larger distances in the Galaxy and indicates that the Sun does not lie in a region which is particularly rich or poor in interstellar plasma. The estimates of the local space density of pulsars made by Taylor and Manchester⁵ and by Davies, Lyne and Seiradakis² were based on assumed mean electron densities of 0.03 and 0.025 cm^{-3} respectively. As the most numerous pulsars in the Galaxy have a very low luminosity, we observe them only within a few hundred parsecs of the Sun. To estimate the local space density of pulsars reliably, it is important to have a good estimate of the local mean electron density near the Sun, or better still a direct measure of distance. It seems therefore that their estimates of the pulsar space density will not be seriously in error due to this assumption and that the rather high pulsar birth rates of between 1 every 4 and 1 every 15 yr required to sustain the derived galactic pulsar population are accordingly still justified. More precise measurements are in progress and will help to determine whether supernovae occur often enough to be the sole birth places of pulsars.

M. J. SALTER
A. G. LYNE
B. ANDERSON

University of Manchester,
Nuffield Radio Astronomy Laboratories,
Jodrell Bank, Macclesfield,
Cheshire, UK

Received 2 May; accepted 4 July 1979.

1. Lyne, A. G. *IAU Symp. No. 60*, 87 (1974).
2. Davies, J. G., Lyne, A. G. & Seiradakis, J. H. *Mon. Not. R. astr. Soc.* **179**, 635 (1977).
3. Anderson, B., Lyne, A. G. & Peckham, R. J. *Nature* **258**, 215 (1975).
4. Taylor, J. H. & Manchester, R. N. *Astr. J.* **80**, 794 (1975).
5. Taylor, J. H. & Manchester, R. N. *Astrophys. J.* **215**, 885 (1977).

On the Pioneer 11 observation of the ring of Jupiter

SHORTLY before closest approach to Jupiter, the imaging experiment on board Voyager 1 detected a ring of particulate matter around the planet as the spacecraft crossed the equatorial plane of Jupiter. The radial position of the ring is at approximately $1.7R_J$ (jovian radii), the total ring thickness is reported to be $\leq 30 \text{ km}$ and the radial extent $\leq 10^4 \text{ km}$ (refs 1, 2). Although this discovery was rather unexpected, the possible existence of a dust belt within the orbit of the innermost satellite, Amalthea, actually had been suspected as a result of the Pioneer 11 measurement of the jovian trapped radiation belt^{3,4}. In this note we briefly outline this interesting observation and evaluate some of the basic properties of this newly discovered ring.

During the Pioneer 11 flyby of Jupiter in December 1974, the minimum distance of the spacecraft to the centre of the planet was as small as $1.6R_J$ and hence the intensity variations of the trapped radiation belt charged particles between the outer magnetosphere and the region around $1.6R_J$ were sampled by energetic particle detectors. As in the earlier Pioneer 10 encounter, absorption structures near the orbits of Ganymede ($L = 15.0$), Europa ($L = 9.4$) and Io ($L = 5.9$) were found in the particle fluxes in certain energy channels^{3,5-7}. Sometimes the

intensity dropped by a factor of 10 in the vicinity of the satellite sweeping corridors. Because of its closer encounter distance, Pioneer 11 enabled a survey to be made of the jovian radiation belt near Amalthea. As well as two absorption features apparently related to Amalthea there are two minima at $L \approx 1.6$ and 1.7 not accounted for by any known satellite³. On one hand, such minima may be simply caused by the anomalous precipitation of the mirroring charged particles into the planetary upper atmosphere at this radial distance as a result of the large quadrupole moment of the jovian magnetic field^{4,8}. On the other hand, the absorption effect of a small satellite or a ring of dust particles can also produce just such structure in the intensity variations of the charged particles^{3,4}. It has been noted, in any case, that even with the anomalous precipitation effect operating in reducing the particle fluxes near the minima in question some particulate matter may still exist and help to sweep up the charged particles in this region⁹. In the light of the Voyager 1 discovery, it seems appropriate to reconsider the Pioneer 11 data in terms of the dust belt interpretation. Because the distances of closest approach to Jupiter for both Voyager spacecraft are more than $4.9R_J$, no information can be obtained for the trapped radiation belt in the vicinity of the ring. This certainly makes the measurements of Pioneer 11 even more interesting. In this letter, we make some estimates (in a manner essentially similar to the treatment discussed by Van Allen¹⁰ for the saturnian rings) of the optical thickness of the ring, number density, as well as the average size of the ring particles. Only proton ($E \geq 80 \text{ MeV}$) data from the UCSD Trapped Radiation Detector instrument described by Fillius³ will be used in the present discussion. This calculation, together with the Voyager 1 imaging result, should provide preliminary information on the new ring of Jupiter.

The satellite absorption effect is determined by two basic factors¹¹⁻¹³: (1) the effective absorption cross-section of the satellite, and (2) the time scale for the charged particles to diffuse across the sweeping corridor of the satellite. Furthermore, the absorption process will be affected by the drift frequency and pitch angle of the particle motion in the dipole field of the planet. If, for a certain kinetic energy, the drift frequency is comparable to the frequency of the relative motion between the satellite and the co-rotating magnetosphere and the two motions are in opposite directions, the resulting 'resonance' would tend to suppress the satellite absorption effect. At the same time, because of the 10° tilt of the magnetic dipole moment with respect to the rotation axis of the planet, mirroring charged particles with equatorial pitch angle larger than 70° would be able to diffuse past the satellite more easily without being absorbed¹¹. In the case of absorption of charged particles by a ring of particulate matter, the pitch angle effect might still be important in determining the degree of absorption; but the drift frequency should not influence the particle absorption since the ring particles are presumably distributed uniformly along the whole orbit. The only energy-dependent factor then would be the stopping range of the charged particles, that is the integrated mass intercepted by the charged particles as they diffuse across the ring plane. For 100-MeV protons, the stopping range is typically a few grammes for matter of various composition (ice, aluminium and so on). Hence, if most of the ring particles have sizes of $>1 \text{ cm}$, just one pass would be enough to absorb the impacting proton; and the survival probability of the energetic protons as they diffuse inward past the ring can be written as

$$P = \exp(-2N_b\tau_\perp) \quad (1)$$

where N_b is the total number of bounces executed by the mirroring protons and $\tau_\perp$ is the optical thickness in the direction perpendicular to the ring plane. The value of $\tau_\perp$ is related to the number density and average diameter (d) of the ring particles, and the thickness of the ring plane ($l_\perp$) by the following expression,

$$\tau_\perp = \pi n d^2 l_\perp / 4 \quad (2)$$

The value of N_b can be approximated by taking the ratio of the pertinent diffusion time scale t_d to the bounce period t_b of the

100-MeV protons. In the Pioneer measurements, the radial diffusion coefficient (D) for 100-MeV proton is estimated to be $\approx 10^{-6} \text{ s}^{-1}$ at the orbit of Io ($L = 5.9$) and 10^{-8} s^{-1} at the orbit of Amalthea ($L = 2.6$) (see ref. 12). If D is proportional to L^3 or L^4 the corresponding value at $L \approx 1.7$ is $\approx 7 \times 10^{-9} \text{ s}^{-1}$. Now, according to the preliminary report^{1,2}, the total radial extent of the ring is $\geq 9 \times 10^3 \text{ km}$ ($\Delta L \geq 0.1$). At the same time, an upper limit for ΔL of about 0.16 can be set from the Pioneer data³. The diffusion time scale can then be approximated as¹⁴:

$$t_d = \Delta L^2 / 4D \quad (3)$$

with $\Delta L \approx 0.1$.

Therefore $t_d \approx 3.6 \times 10^5 \text{ s}$. The bounce period of mirroring charged particles is given by¹⁵

$$t_b \approx 5.2 r_0 / V \quad (4)$$

where r_0 is the equatorial radial position of the mirroring charged particle and V its velocity. For $r_0 = 1.7 R_J$ and $V = 1.2 \times 10^{10} \text{ cm s}^{-1}$ we have $t_b \approx 5.2 \text{ s}$; and hence $N_b \approx 7 \times 10^4$. Judging from the depth of the minima in the proton intensity as measured by the UCSD experiment we can put $P = 0.1$ (that is, the flux is reduced by a factor of 10 across the absorption structure). From this we can immediately obtain $N_b \tau_{\perp} \approx 2.3$ and therefore $\tau_{\perp} = 6.2 \times 10^{-6}$ for the absorbing particles. Such a small optical thickness may be compared with $\tau_{\perp} \approx 1$ for the B ring of Saturn¹⁶ and a value between 10^{-7} and 10^{-2} for the D' ring suspected to exist beyond the outer edge of the A ring^{17,18}. [The result that the $\tau_{\perp}$ value of the ring of Jupiter should be a factor of 6×10^{-6} smaller than that of the saturnian B ring is in good agreement with the most recent IR detection of Jupiter's ring¹⁹. In this observation it is concluded that the number of particles in Jupiter's ring should be of the order of 10^5 less than that in Saturn's if both ring systems have similar kinds of particles and the viewing geometries are about the same.]

If the ring plane thickness $l_{\perp}$ is known we can derive the product nd^2 using equation (2). No exact value of $l_{\perp}$ is yet known—only that it is less than 30 km (refs 1, 2). Considering a thickness of $\sim 1 \text{ km}$ estimated for the saturnian rings²⁰ it may be reasonable to adopt $l_{\perp} \approx 1 \text{ km}$ for the ring of Jupiter as a working number. The actual thickness is not likely to be larger or smaller by a factor of 10. Substitution of this value into equation (2) gives $nd^2 \approx 8 \times 10^{-11} \text{ cm}^{-1}$; that is

$$n \approx 8 \times 10^{-11} / d^2 \text{ cm}^3 \quad (5)$$

With the total volume of the ring given as $U = 2\pi r_0 \Delta r l_{\perp}$ (thus $U \approx 6 \times 10^{24} \text{ cm}^3$ if $r_0 = 1.7 R_J$, $\Delta r = 0.1 R_J$ and $l_{\perp} = 1 \text{ km}$) and the density of the ring particles is 2 g cm^{-3} , the total mass of the ring can be expressed as

$$M \approx 5 \times 10^{14} d \text{ grammes} \quad (6)$$

Thus $n \approx 8 \times 10^{-11} \text{ cm}^{-3}$ and $M \approx 5 \times 10^{14} \text{ g}$ if $d \approx 1 \text{ cm}$, and $n \approx 8 \times 10^{-21} \text{ cm}^{-3}$ and $M \approx 5 \times 10^{19} \text{ g}$ if $d \approx 1 \text{ km}$.

From simple geometrical considerations we find the maximum optical path ($l_{\parallel}$) along the ring plane as it is observed edge-on to be about $1.1 R_J$, therefore the corresponding optical thickness ($\tau_{\parallel}$) can be approximated as $\tau_{\perp} \times (l_{\parallel} / l_{\perp})$, and this gives $\tau_{\parallel} \approx 0.5$. If the ring particles have a certain size distribution, the bodies with diameter much less than 1 cm may contribute significantly to a total light reflecting area ($A \approx 6 \times 10^{14} \text{ cm}^2$ if $\tau_{\perp} \approx 6 \times 10^{-6}$) even though their absorption power is rather limited (since most of the mass is likely to be concentrated in the large particles). In this manner, the effective optical thickness may be increased by one or two orders of magnitude—but probably not by a factor much larger—considering the absence of the ring from ground-based observations in the past. If we argue that the ring particles are as old as the Solar System, their size must be large enough to resist orbit decay caused by the Poynting–Robertson effect. Following the similar treatment for the saturnian ring particles²¹, the minimum diameter can be estimated to be $\approx 1 \text{ m}$, at least, for particles carrying most of the mass. A determination of the size distribution could possibly be achieved by analysing the absorption effect on electrons and protons of different energies (since they have different stopping ranges). Before this is attempted, little can be said.

In this connection, we would like to mention that detailed examination of the Pioneer 11 data with a small time resolution may also provide vital information on the radial structure of the ring. For example, the radial distance of the 2:1 two-body resonance with Amalthea is $1.61 R_J$. It would be interesting to see if there is energetic particle signature along the magnetic flux tube in this vicinity indicating the existence of a gap reminiscent of the Cassini's division in the saturnian rings. Or, the orbital resonance may actually define the mass concentration of the ring as previously suggested for the narrow rings of Uranus (but for a different type of orbital resonances) (see refs 22–24). The Pioneer 11 energetic charged particle data might then provide information on the dynamical models of ring formation as well.

Thus, based on the Pioneer 11 UCSD proton ($E \geq 80 \text{ MeV}$) measurement of the absorption structures in the vicinity of the ring of Jupiter, the optical thickness of the absorbing ring particles in the direction perpendicular to the ring plane is estimated to be about 6×10^{-6} , and the corresponding value as the ring is seen edge-on is approximately 0.5. If the ring is as old as the Solar System, most of the mass should be concentrated in particles with diameter larger than 1 m and the total mass of the ring would be larger than $5 \times 10^{16} \text{ g}$. Comparison with the Voyager imaging data may narrow down the uncertainty. But only further analysis of the energetic charged particle data can give us more insight about the size distribution and radial structure of the ring.

I thank Dr Walker Fillius for very useful discussions since his finding of the 'anomalous' absorption minima inside the orbit of Amalthea. I also thank Professor Ian Axford for a discussion on the L -dependence of the radial diffusion coefficient.

W.-H. IP

Max-Planck Institut für Aeronomie,
D-3411 Katlenburg-Lindau 3, FRG

Received 20 March; accepted 27 June 1979.

- Smith, B. A. *et al. Science* **B 204**, 951–973 (1979).
- Beatty, K. J. *Sky Telescope* **57**, 423 (1979).
- Fillius, W. in *Jupiter* (ed. Gehrels, T.) 896 (University of Arizona Press, Tucson 1976).
- Acuna, M. H. & Ness, N. F. *J. geophys. Res.* **81**, 2917 (1976).
- Van Allen, J. A. in *Jupiter* (ed. Gehrels, T.) 928 (University of Arizona Press, Tucson, 1976).
- McDonald, F. B. & Trainor, J. H. in *Jupiter* (ed. Gehrels, T.) 961 (University of Arizona Press, Tucson, 1976).
- Simpson, J. A. and McKibben, R. B. in *Jupiter* (ed. Gehrels, T.) 738 (University of Arizona Press, Tucson, 1976).
- Roederer, J. G., Acuna, M. H. & Ness, N. F. *J. geophys. Res.* **82**, 5187 (1977).
- Ip, W.-H. *Space Sci. Rev.* (in the press).
- Van Allen, J. A. *Highlights of Astronomy* **4**, 195 (1977).
- Mead, G. D. & Hess, W. N. *J. geophys. Res.* **78**, 2793 (1973).
- Mogro-Campero, A. in *Jupiter* (ed. Gehrels, T.) 1190 (University of Arizona, Tucson, 1976).
- Thomsen, M. F. *Rev. Geophys. Space Phys.* **17**, 369 (1979).
- Mogro-Campero, A. & Fillius, W. *J. geophys. Res.* **81**, 1289 (1976).
- Roederer, J. G. *Dynamics of Geomagnetically Trapped Radiation* (Springer, Heidelberg, 1970).
- Bobrov, M. S. in *Surfaces and Interiors of Planets and Satellites*, (ed. Dollfus, A.) 377 (Academic, New York, 1970).
- Ferrin, I. *Icarus* **22**, 159 (1974).
- Smith, B. A., Cook, A. F., Feibelman, R. F. & Beebe, R. F. (1975). *Icarus*, **25**, 466.
- Becklin, E. E. & Wynn-Williams, C. G. *Nature* **279**, 400 (1979).
- Fountain, J. W. & Larson, S. M. *Icarus* **36**, 92 (1978).
- Pollack, J. B. *Space Sci. Rev.* **18**, 3 (1975).
- Dermott, S. F. & Gold, T. *Nature* **267**, 590 (1977).
- Ip, W.-H. *Nature* **272**, 802 (1978).
- Steingmann, G. A. *Nature* **274**, 454 (1978).

Compositional differences between Arctic aerosol and snow

UNIQUE information on trace elements in polar atmospheres is available through records of deposition in snow and ice. Proper interpretation of these data requires a knowledge of the nontrivial chemical relationship between the deposition and its parent aerosol. Because several complex processes determine the trace-element content of precipitation, polar ice and snow cannot be considered *a priori* to have the same composition as polar aerosol¹. In most of the current literature, such an assumption is usually made. Within the past year or so, the first tests of the long-term relationship between aerosol and deposition have been (inadvertently) produced for the Antarctic and

Table 1 Elemental concentrations and crustal enrichment factors in Barrow aerosol and snow

Element	Aerosol*		Snow†	
	Concentration (ng m ⁻³)	EF _{crust} ‡	Concentration (µg l ⁻¹)	EF _{crust} ‡
Na	770	74	144	15.9
Mg	160	21	41	6.1
Al	30	1.0	26	1.0
Ca	55§	4.1	100	8.6
V	0.65	13.0	0.086	2.0
Mn	1.12	3.2	0.79	2.6
Zn	14.8§	580	1.16	52
Cd	0.37§	5,000	0.047	730
Hg	<0.4§	<14,000	0.011§	430

*Mean of December 1976–April 1977 and December 1977–April 1978 at Barrow, Alaska.

†From ref. 5, 28–224 km south of Barrow, Alaska, 15 February 1974. Na, Ca from 28- and 84-km points; Cd from all except 28-km point; other elements from all except 56-km point.

‡For element X, $EF_{crust} = (X/Al)_{aerosol \text{ or snow}} / (X/Al)_{rock}$, where the rock composition is from ref. 10.

§Estimated from incomplete series of measurements; uncertainties a factor of 2.

the Arctic. The results are different for each region, and illustrate the great caution that must be exercised in this entire field. At the South Pole, snow from 1973–74² compares very well with aerosol from 1974–75³—both media give the same impression of the environment^{2,4}. In the Arctic, on the other hand, opposite conclusions about the atmosphere have been drawn from snow and aerosol, and this problem is discussed here.

Weiss *et al.*⁵ have used crustal enrichment factors of trace elements in recent Barrow, Alaska, snow and pre-1900 Greenland ice to conclude that (1) the 'lithospheric' elements Al, V, and Mn are not enriched at Barrow and Milcent (Greenland), and hence are derived from a (natural) 'windborne lithospheric component'; (2) the enrichments of Na, Zn and Cd are similar at Barrow, Milcent, and other Greenland sites (which implies that these elements are natural in origin at Barrow); (3) other enriched elements such as S, Se, Hg, Cd, As and Sb are naturally enriched in pre-1900 Greenland deposits. The overall impression given by Weiss *et al.*⁵ is of the natural character of polar deposition in both Greenland before 1900 and Barrow today. Weiss *et al.* refer interchangeably to 'atmosphere' and 'aerosol' 'snow', 'ice', and 'deposit', implying that this natural character also holds for the polar atmosphere.

In contrast, we have recently concluded from chemical analysis of two years' samples of the Barrow aerosol that it is strongly pollution-derived, particularly during winter, to the extent that abundances of even 'lithospheric' elements can be affected⁶. This deduction is based on actual V/Al and Mn/Al ratios, which are more characteristic of mid-latitude pollution than of crust. It is supported by ancillary evidence such as abundant grey, sooty carbon and high concentrations of sulphate (1–5 µg m⁻³) in the aerosol⁶ (both of which imply pollution), and high concentrations of ²¹⁰Pb in the aerosol^{6,7} (which indicate that it is continental in origin). Tests at the surface (spring 1979) and aloft⁸ have shown that this pollution does not originate on the North Slope; it is associated with air coming to Alaska from farther north. In analogy to the similarity of snows from various parts of the Arctic reported by Weiss *et al.*, we find a similar (pollution) aerosol at the widely dispersed Arctic sites of Spitsbergen, northern Norway, and northern Greenland during winter⁶.

We therefore have two lines of evidence (aerosol and snow) which have arisen independently, are both solid and defensible, but offer divergent pictures of the Arctic environment. How can they be reconciled? Part of the difference for Barrow can be explained by recalculating snow–crust enrichment factors with Al rather than Mn as crustal reference element. Weiss *et al.* used Mn because it was common to all their snow samples. In general, however, Mn is enriched by factors of 3–10 in urban

and rural atmospheres⁹, presumably by pollution sources, and therefore should not be used as a crustal reference element wherever it can be avoided. (Of course, without knowing how much the Barrow aerosol is affected by pollution, it would seem reasonable to use Mn.) Concentrations and revised enrichment factors for the Barrow winter aerosol and the near-Barrow snows are shown in Table 1. These calculations, the basis for which can be justified from many aerosol samples⁹, show that both V and Mn have modest but definite enrichments in the near-Barrow snows. But this recalculation removes only part of the discrepancy between Barrow aerosol and snow—V and Mn are still several times more enriched in the aerosol than in the snow, and all other elements except Ca have greater enrichments in aerosol than in snow. For Na, Mg and Ca, this difference may be marine-related, because the aerosol-sampling site was closer to the sea than were the snow-sampling sites. The other elements, however, are not marine in origin; their different enrichments must have another explanation.

We believe that the broad discrepancy between Arctic aerosol and Arctic snow is a genuine feature of the present Arctic environment. Possible explanations for it include: (1) formation of snow from a parent aerosol aloft whose composition differs from the surface aerosol; and (2) aerosol–snow fractionation during or after snowfall. We favour the second explanation for two reasons: first, we see little reason that the composition of the aerosol in such a remote, source-free area should change markedly with altitude. Second, formation of snow crystals by nucleation on clay minerals, such as are found at the centre of many or most polar snow crystals,^{11,12} could fractionate snow chemically relative to aerosol. A central clay nucleus which contained 90% of the foreign matter if the ice crystal (estimated by M. Kumai, personal communication) would decrease enrichments in snow by the factor of 10 that is observed. We stress, however, that aerosol–snow fractionation in the Arctic is only a hypothesis.

We emphasise that whatever the eventual explanation, aerosol–snow differences in the Arctic imply, care is needed in interpreting ice and snow data atmospherically. Modern snows may not be automatically equated with modern aerosols; historical trends of deposition may not be immediately equated with historical trends of parent aerosol. Until thorough studies of the factors controlling the elemental composition of polar snow and ice are undertaken, these uncertainties will remain.

This work was supported by the Office of Naval Research (contract N00014-76-C-0435). Barrow aerosol samples were collected cooperatively with the NOAA GMCC program. We thank the staff of the GMCC Clean Air Observatory and the Naval Arctic Research Laboratory in Barrow, Alaska, for assistance. Aerosol samples from Spitsbergen and northern Norway were provided by the Norwegian Institute for Air Research; samples from Thule were provided by the Danish Air Pollution Laboratory and the Danish Meteorological Institute. Samples were analysed chemically at the Rhode Island Nuclear Science Center with assistance from R. D. Borys and T. J. Conway.

KENNETH A. RAHN
RICHARD J. MCCAFFREY

Graduate School of Oceanography,
University of Rhode Island,
Kingston, Rhode Island 02881

Received 29 May; accepted 29 June 1979.

1. Junge, C. E. in *Proc. Symp. Isotopes and Impurities in Snow and Ice* IAHS Publ. No. 118, 63–77 (1977).
2. Boudron, C. & Lorius, C. *Nature* **277**, 551 (1979).
3. Maenhaut, W. & Zoller, W. H. *J. radioanalyt. Chem.* **37**, 637–650 (1977).
4. Zoller, W. H., Cunningham, W. C. & McGregor, C. *Ant. J. U.S.* **13**, 187–188 (1978).
5. Weiss, H. V., Herron, M. M. & Langway, C. C. Jr *Nature* **274**, 352–353 (1978).
6. Rahn, K. A. & McCaffrey, R. J. in *Proc. Conf. on Aerosols: Anthropogenic and Natural Sources and Transport* (New York Academy of Sciences, in the press).
7. *Environmental Quarterly, Environmental Measurements Laboratory Rep. EML-344*, UC-11, Appendix (US Department of Energy, New York, 1978).
8. Rahn, K. A., Borys, R. D. & Shaw, G. E. *Nature* **268**, 713–715 (1977).
9. Rahn, K. A. *The Chemical Composition of the Atmospheric Aerosol* (University of Rhode Island, 1976).
10. Mason, B. *Principles of Geochemistry*, 3rd edn (Wiley, New York, 1966).
11. Kumai, M. *Geophys. pura appl.* **36**, 169–191 (1957); *J. Met.* **18**, 139–150 (1961); *J. Atmos. Sci.* **33**, 833–841 (1976).
12. Kumai, M. & Francis, K. E. *J. Atmos. Sci.* **19**, 474–481 (1962).

The 20-yr cycle in Greenland ice core records

CLIMATIC oscillations with a period of ~ 20 yr have been reported in North American winter temperature records^{1,2} midwestern US drought indicators (ref. 3 and J. M. Mitchell, Jr personal communication), dendrochronological data⁴, and sea-surface temperature anomalies⁵. These oscillations have mostly been attributed to double sunspot cycles, which seem to occur at about 22-yr intervals. Other possible causes include an 18.6-yr modulation of the diurnal and semidiurnal lunar tidal forces⁵. Critics^{6,7} of these observations argue that such oscillations may only be an artefact of a red noise record, and are not uniformly present in a wide range of climatic data. It has also been argued that even if the oscillations are significant, their small amplitude renders them relatively unsuitable for climatic prediction¹. However, if such oscillations prove to be real, identifying their nature and possible causes, even though their amplitude is small, can enhance our understanding of climatic change. To gain more insight into such 20-yr oscillations, we have examined characteristic oscillations in Greenland ice core oxygen isotope records. The results show that the strongest spectral peak occurs at a period of 20 ± 0.5 yr with a statistically significant amplitude. This period is essentially the same as that found in eastern North American winter temperatures¹. Moreover this 20-yr oscillation is found to be coherent in phase with the dominant variation in the Sun's motion about the centre of mass of the Solar System. This motion has been hypothesised⁸ to be closely linked with sunspots. On the basis of these observations we speculate here that a 20-yr oscillation may be a more fundamental solar and/or climatic oscillation than the nominal 22-yr cycle often mentioned.

Because dating procedures⁹ are highly accurate, climatic time series obtained from Greenland ice core records are especially suited for detailed spectral analysis. From recent cores several well-dated coincident oxygen isotope records spanning 728 yr (1244–1971) have been obtained. These data supply an adequate-length time series for resolving frequencies accurately, while still yielding a typical sampling interval of < 0.2 yr. These records do, however, contain significant amounts of spatially incoherent non-climatic noise due to depositional irregularities. In an analysis of several shallow cores separated by < 60 km, Reeh *et al.*¹⁰ found this non-climatic component accounted for about 50% of the variance of the isotopic records and generally had a 'red' spectral content. This 50% figure has also been verified using deeper (100 m) cores (N. Reeh, personal communication). Because of this depositional noise, the significance of spatially coherent oscillations in the records can be substantially enhanced by averaging several records at different locations.

Although not critical for this study, it should be noted that the precise meteorological significance of the isotopic records is not determined. In Greenland, correlations suggest that the isotopic records indicate atmospheric temperature variations¹¹. Comparison of one of these records¹² with temporally smoothed Icelandic and central England temperatures supports this interpretation. However, the isotopic values are possibly indicative of shifts of storm tracks, which in turn tend to modify the source of the precipitation.

For analysis here, coincident portions of three cores were used: Crete [71.12° N, 37.32° W], Milcent [70.31° N, 44.58° W] and Dye 3 [65.18° N, 43.82° W]. All three of these cores were obtained as part of the International Greenland Ice Sheet Program (GISP). To reduce the depositional noise, oxygen isotopic records from these three cores were averaged to form a combined time series for portions of the analysis.

Of the core sites Milcent is on the western side of the ice divide, Dye 3 on the eastern side and Crete effectively on the divide. Primarily because of this, Dye 3 and Milcent tend to

receive precipitation from different directions, whereas Crete probably receives precipitation from either direction¹³. Also, Dye 3 is further south, and hence the core contains more melt features which yield some measure of summer temperature¹⁴.

To analyse these isotopic records spectrally both conventional lagged product¹⁵ and maximum entropy¹⁶ (ME) spectral analysis methods were used. Conventional techniques allow statistical confidence levels for the spectral peaks to be estimated, while the ME techniques aid in precise determination of frequencies and power. For the maximum entropy analysis, the Burg algorithm as formulated by Andersen¹⁷ was used to estimate a minimum delay autocorrelation function from which the ME spectrum was calculated. In all cases 2 yr non-overlapping averages were taken before analysis in order to display the low frequencies better and to reduce the noise level. The use of such averages was supported by analyses of the whole 1 yr averaged combined record which yielded low frequency

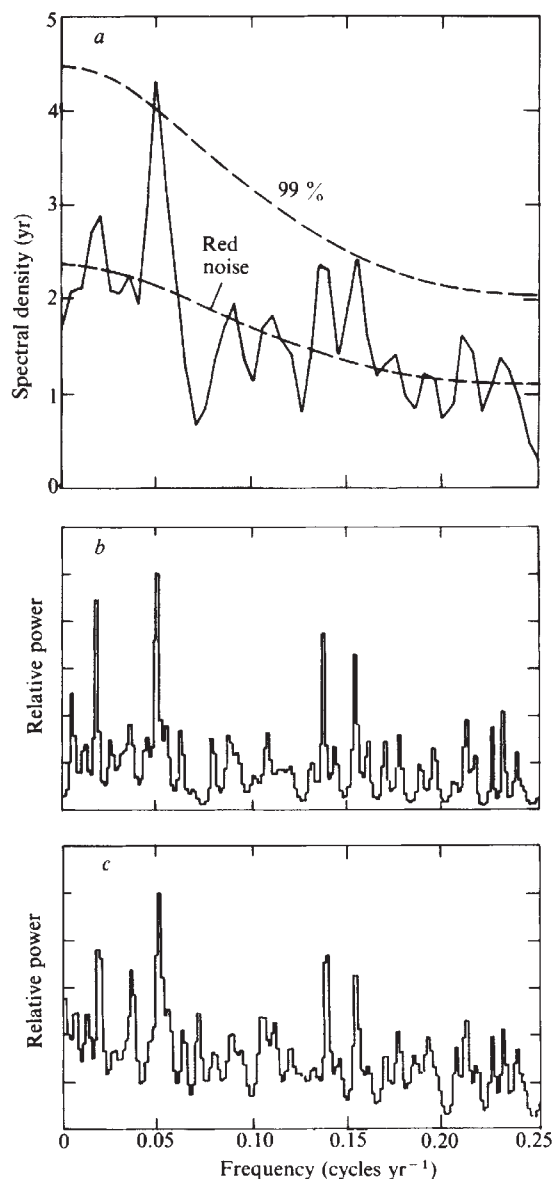


Fig. 1 Spectral densities of ice core oxygen isotope records. Two-year averaged data were used for the analysis. *a*, *b*, The conventional and ME spectra of the combined isotopic record formed by combining the three core records from 1244 to 1971. *c*, The average spectrum of the three records taken separately. In *a* 50 lags were used to analyse a 364-point time series of 2-yr averages, together with a Hamming Spectral Window. Twenty degrees of freedom were used in estimating confidence limits. This estimate was based on the expression $2.7 T/M$ (ref. 15) where T is the record length and M is the auto covariance length.

Table 1 Periods and relative power of ME spectra of combined Crete, Milcent, Dye 3 oxygen isotopic ratio record

Whole record (1244–1971) (80 filter wts)		Bottom half of record (1244–1607) (40 filter wts)		Top half of record (1608–1971) (40 filter wts)	
Period	Relative power	Period	Relative power	Period	Relative power
20.00	8.17	6.47	16.97	20.36	17.28
54.41	5.88	28.35	12.73	7.35	12.51
6.53	5.59	56.67	10.00	51.27	11.09
7.31	5.00	19.73	9.75	144.20	7.52
183.70	4.27	5.28	6.36	9.32	6.87

2-yr averages sampled every 2 yr were used in the analysis.

(≤ 0.25 cycles yr^{-1}) results almost identical to those obtained from 2-yr averaged data.

Figure 1a shows the lagged product spectrum of the combined isotopic record. Figure 1b is the ME spectrum of the combined isotopic record, and Fig. 1c is an average ME spectrum obtained by finding a common unit prediction filter for all three records—a straightforward extension of the Burg algorithm. (This average spectrum is essentially the sum of the individual time series spectra.) The primary difference between the 'combined' spectrum and the 'average' spectrum is that in the 'combined' spectrum cycles with phase differences will cancel, whereas in the 'average' spectrum phase differences from record to record do not matter. Note that both ME spectra have similar dominant peaks with the 'combined' spectrum having less noise. This indicates a phase coherence between the dominant oscillations in the different records. This phase coherence was also verified directly for the 20-yr cycle by direct correlation of bandpassed individual records.

The dominant feature in these spectra is a pronounced peak with about a 20-yr wavelength. As demonstrated in Fig. 1a, the 20-yr peak is statistically significant at the 99% level when compared with red noise. Based on this spectrum it seems unlikely that this peak is simply an artefact of simple 'persistence' in the climate (which in the simplest case is a first order Markov process with a red noise spectrum¹⁸).

To determine more precisely the frequency of this peak, the analytical integration procedure of Johnsen and Andersen¹⁹ was used. In this method, the analytical expression for the maximum entropy spectrum is integrated using residue theory. This procedure gives a precise value for both the power and frequency of a given peak. Table 1 lists the first five peaks (ordered by their power) for the average isotopic record. Also listed are the first five peaks in the bottom and top halves of the record. While there is some scatter in these results we feel they conservatively place the period of the 20 yr oscillation at 20.0 ± 0.5 yr. This result was also verified by higher resolution conventional lagged product spectra with a frequency resolution of 0.0025 cycles yr^{-1} .

This period agrees with a 20-yr period found in eastern North American January temperatures by Mock and Hibler¹. It is, however, in contrast to a 23-yr cycle found by Bain²⁰ in central England temperatures and a 22-yr spectral peak in Bristlecone pine data found by Epstein and Yapp⁴. (Epstein and Yapp⁴, however, used 10-yr averaged data, which could effectively filter out a 20-yr cycle. For example, taking 10-yr averages of the combined isotopic record reduced the 20-yr cycle amplitude and shifted it to 21 yr.) This period also contrasts somewhat with a 20–22 yr drought cycle in the Midwest suggested by Roberts and Olsen³ and found in tree-ring drought data by Mitchell, Stockton and Meko (personal communication).

One of the main reasons that researchers favour a 22-yr cycle in climatic records has been the belief in possible weather links to a hypothesised 22-yr double solar cycle. In principle, this cycle arises because of different magnetic polarities²¹ of sunspot cycles which, in observations over the past 90 yr, have

been found to reverse in alternate cycles nominally occurring at 11 yr intervals. It also appears that magnetic activity is consistently different in adjacent cycles²¹. Pittcock⁷ has argued that no link between solar cycles and climatic change has been established, and a 20-yr rather than a 22-yr cycle could support his scepticism. By this view the 20-yr cycle in the ice core data could be simply a resonance in the climate system. However, examination of available sunspot data indicates that it is difficult to ascertain the precise period of the 'double solar cycle'. Consequently, an alternative hypothesis which, in our judgement, cannot be ruled out is that a double solar cycle is occurring at 20 yr rather than 22 yr. Under this hypothesis the relatively irregular sunspots could be viewed as a stochastic manifestation of some more regular 20-yr cycle. Certainly such a coincidence would not prove a link between solar variations and the ice core records, but it would at least provide a possible forcing at the correct frequency.

Although he performed no spectral analyses, Jose⁸ indirectly advanced the hypothesis of a 20-yr solar oscillation in a detailed calculation of the Sun's position relative to the Solar System centre of mass. In particular, Jose compared various solar orbital parameters (which typically vary with a 20-yr period) with sunspot data. He found that if, in early sunspot records without polarity observations, the polarities of adjacent sunspot cycles were in special cases not reversed, then excellent correlation with the Sun's motion could be obtained. Note that such a deterministically forced theory of sunspots, while not necessarily in agreement in period, is consistent with a recent suggestion by Dicke²² that the sunspots are irregular manifestations of a more regular phenomenon.

To compare the period and phase of the solar oscillation at 20 yr with the ice core oscillation, we have made a simple (but less precise than Jose) calculation of the Sun's orbital parameters. We assume, as a first approximation, that the planetary orbits are circular, and in the same plane, then by conservation of momentum (that is $\sum_i m_i \dot{x}_i = J$) the position of the Sun relative to the centre of mass of the Solar System may be easily estimated. Using planetary data from Allen²³, we calculated various orbital parameters over the same 728-yr period spanned by the ice core data. Figure 2 shows a ME spectrum of the rotation rate $[(\ddot{x}\ddot{y} - \dot{x}\dot{y})/(\dot{x}^2 + \dot{y}^2)]$ of the centre of the Sun about its instantaneous centre of curvature. This parameter is chosen because it would be a portion of the effective 'coriolis' parameter seen by any rotating turbulence in the convective zone of the Sun. (Parameters such as the Sun's velocity do, however, yield similar spectra.) The spectrum is strongly dominated by a 19.8-yr peak. This peak physically arises due to the synodic period of Jupiter and Saturn, being ~ 19.85 yr. There is also a secondary weaker peak at ~ 11.9 yr which corresponds closely to the 11.86 yr orbital period of Jupiter.

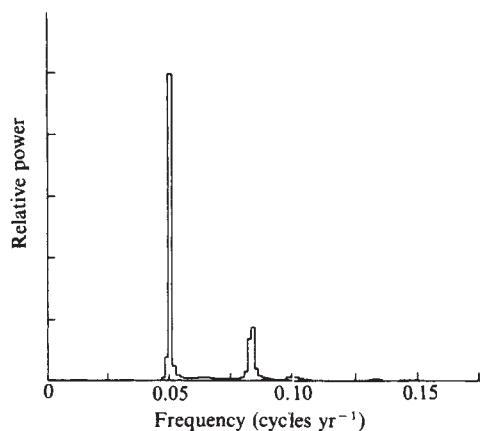


Fig. 2 Spectrum of rotation rate of the centre of the Sun about its instantaneous centre of curvature (relative to a fixed coordinate system at the centre of mass of the Solar System).

Table 2 Periods and relative power of the sunspot record (1699–1971) using various polarities

Single cycle (polarity ignored)		Double cycle (polarity reversed in adjacent cycles)		Double cycle using Jose's polarities	
Period	Relative power	Period	Relative power	Period	Relative power
10.93	33.12	21.99	70.22	20.12	55.82
9.99	20.22	18.37	12.21	24.24	18.76
93.74	11.68	26.39	2.89	30.73	6.75
12.15	7.23	15.87	2.29	13.94	3.81
50.37	6.61	13.71	1.96	16.11	3.76

Annual mean values together with 100 filter wts were used in the analysis.

To establish the phase drift of the isotopic cycle relative to the Sun's 'coriolis' variations, we correlated a narrow frequency band of the isotopic record with the solar motion. To construct this bandpassed record the procedure of Ulrych *et al.*²⁴ was used. In particular the record was first extended forwards and backwards using ME, and then bandpass filtered, using a filter²⁵ with pass band from a period of 21.05 yr to 19.05 yr. Since the ratio of this band width to the Nyquist frequency is about 0.02, we estimate the degrees of freedom to be about 1 per 100 yr. In both the bottom and top halves of the record the best correlation occurred at a 180° phase shift (within 2 yr), indicating a good phase coherence. While this phase coherence could be a statistical fluke, it suggests a deterministic forcing rather than an internal stochastic oscillation.

To demonstrate the effect of Jose's hypothesised polarity on the sunspot spectra, we used the Johnsen-Andersen¹⁹ spectral integration technique to determine the first five peaks in the various sunspot records (see Table 2). As can be seen, using Jose's polarities the inverted cycle yields a sharp 20-yr peak. Note also that in addition to an 11-yr cycle there is a strong 10-yr cycle in the normal (polarity ignored) sunspot record²¹.

Studies by Eddy²⁶ have suggested that the occurrence of sunspots may be quite irregular, with the sunspots even disappearing for several decades. This fact in conjunction with the uncertainty in periods exhibited by Table 2 suggests that it is probably not possible to establish whether a 'double solar cycle' is occurring at 20 or 22 yr from available sunspot records alone.

In conclusion, the results presented here demonstrate a dominant oscillation in Greenland ice core isotopic records at a period of about 20 yr. While we know of no causal link, such a period also coincides with the Sun's motion relative to the centre of mass of the Solar System. This motion has been suggested to correlate with sunspot variations. On the basis of these observations we speculate that a 20-yr oscillation may be a more fundamental climatic variation than the nominal 22-yr 'solar cycle'.

We thank Dr N. Reeh for discussions on ice core spectra, Dr S. Mock for comments on sunspot spectra, Dr N. Andersen for suggesting a simple way to calculate solar motions, Dr W. St. Lawrence for helpful comments, and Professor W. Dansgaard for encouragement. This work was supported by the DA and the NSF.

W. D. HIBLER III

US Army Cold Regions Research
and Engineering Laboratory,
Hanover, New Hampshire 03755

S. J. JOHNSEN

Geophysical Isotope Laboratory,
University of Copenhagen,
Hørlæsgade 6,
DK-2200 Copenhagen, Denmark

Received 23 April; accepted 29 June 1979.

- Mock, S. J. & Hibler III, W. D. *Nature* **261**, 484–486 (1976).
- Spar, J. & Mayer, J. K. *Weatherwise* **26**, 128–130 (1973).
- Roberts W. O. & Olson, R. H. *Nature* **254**, 380 (1975).
- Epstein, S. & Yapp, C. J. *Earth planet. Sci. Lett.* **30**, 252–261 (1976).
- Loder, J. W. & Garrett, C. J. *geophys. Res.* **83**, 1967–1970 (1978).
- Monin, A. S. & Vulis, I. L. *Tellus* **23**, 337–345 (1971).
- Pitcock, A. B. *Rev. Geophys. Space Phys.* **16**, 400–420 (1978).
- Jose, P. D. *Astr. J.* **70**, 193–200 (1965).
- Hammer, C. U. *et al. J. Glaciol.* **20**, 3–26 (1978).
- Reeh, N., Clausen, H. B., Gundestrup, N., Johnsen, S. J. & Stauffer, B. *Proc. Symp. Isotopes and Impurities in Snow and Ice* IAHS Publ. No. 118, 117–181 (1977).
- Dansgaard, W., Johnsen, S. J., Clausen, H. B. & Gundestrup, N. *Meddr. Grønland* **197** (1973).
- Dansgaard, W. *et al. Nature* **255**, 24–28 (1975).
- Reeh, N. *et al. J. Glaciol.* **20**, 27–30 (1978).
- Hibler, W. D. III & Langway, C. C. Jr. *Polar Oceans* (ed Dunbar, M. J.) 589–602 (A.I.N.A. 1977).
- Jenkins, G. M. & Watts, D. G. *Spectral Analysis and its Applications* (Holden-Day, San Francisco, 1969).
- Burg, J. P. thesis, Stanford Univ. (1975).
- Andersen, N. *Geophysics* **39**, 69–72 (1974).
- Gilman, D. L., Fuglister, F. J. & Mitchell, J. M. Jr. *J. Atmos. Sci.* **20**, 182–184 (1963).
- Johnsen, S. J. & Andersen, N. *Geophysics* **43**, 681–690 (1978).
- Bain, W. C. Q. *Jl. met. Soc.* **102**, 464–466 (1976).
- Chernosky, E. J. *J. geophys. Res.* **71**, 965–974 (1966).
- Dicke, R. H. *Nature* **276**, 676–680 (1978).
- Allen, C. W. *Astrophysical Quantities* (University of London Press, 1955).
- Ulrych, T. J., Smylie, D. E., Jensen, O. G. & Clarke, G. K. C. *J. geophys. Res.* **78**, 4959–4964 (1973).
- Hibler, W. D. III, *J. geophys. Res.* **77**, 7190–7195 (1972).
- Eddy, J. A. *Science* **192**, 1189–1202 (1976).

Combined structural and chemical analysis of 3,800-Myr-old microfossils

CELL-like inclusions detected in the cherty layers of a quartzite, which is part of the Isua series in South-west Greenland, consist of biological materials, according to analyses by Raman laser molecular microprobe. The available radiometric data place the age of the sequence at around 3,800 Myr¹. Thin sections of our specimens were taken in their primary positions within the rock matrix. The material used was compact and unweathered. No maceration, etching, impregnation or other methods were applied which might have produced artefacts.

About 100 well preserved specimens noted in the material were sealed within the quartz grains or in the silica cement in a three-dimensional condition (Fig. 1d, f). This sealing apparently results from a syngenetic permineralisation caused by colloidal silica, which subsequently crystallised to form chalcedony. Our observations suggest that these specimens are primary constituents of the sediment, and are not contaminants from a later time.

The fossils occur as individual unicells (Fig. 1c), filaments (Fig. 1a, b) or cell colonies (Fig. 1e). Cells and cell families are usually surrounded by multilaminar sheaths which show a characteristic laminar structure. All specimens observed apparently belong to the same kind of organism, named *Isua-sphaera*².

The individual cells are more or less ellipsoid in shape. The mature cells range between 20 and 40 µm in diameter. The cell encloses a more or less globular hollow which is partly filled with organic matter (p in Fig. 1e). Apparently, this filling is a remnant of the former protoplasm which has been degraded during fossilisation. A vacuole is often contained in the cell lumen (v in Fig. 1c).

The mode of reproduction indicated resembles the kind of budding known in recent yeasts. Figure 2 shows the characteristic stages of reproduction frequently observed in the material.

The best way of confirming the biogenicity of a fossil residue is by relating morphology and chemistry. In practice, the chemical proof becomes a problem: most of the conventional analytical techniques require powdering and leaching of the rock samples. Consequently the results obtained are representative of the whole-rock bitumen and not specifically the enclosed microfossils.

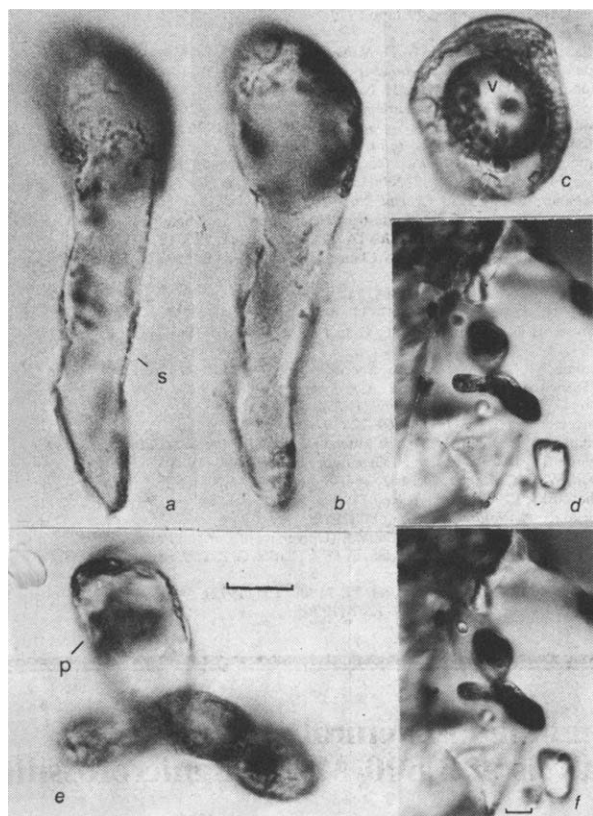


Fig. 1 *Isuasphaera isua* Pflug, 1978 Isua quartzite Greenland 3,800 Myr *a, b*, Filament: *s*, sheath; *c*, single cell: *v*, vacuole; *d-f*, ramified colony: *p*, organic cell filling; *d, f* are two foci, showing the specimen in its primary position within a quartz grain (bright). Magnifications: *a-c, e* $\times 900$; *d-f*, $\times 290$. Scale bar in *e* and *f*, 10 μm .

Individual and non-destructive analysis of microfossils in rocks has proved possible with the Raman microprobe analyser Mole³. This is a new physical method of microanalysis, based on the vibration spectra characteristic of polyatomic structures. A focused laser beam excites the sample. The light diffused by the Raman effect is used to identify and locate the molecular constituents present in the sample⁴.

For the present work only one of the possible modes of the Mole was used: the monochannel spectrometer Raman mode. A light-optical microscope investigates the interesting structures in the sample and where the laser beam should be directed. The measuring spot can be restricted to $\sim 1 \mu\text{m}$ in diameter. The diffused light is collected and analysed by a double monochromator equipped with holographic gratings. It is then received by a photomultiplier, amplified and recorded on paper as a function of wavenumber. The light used for excitation was the green line (5,145 Å) of an ionised argon laser.

Unmounted thin sections, $\sim 40\text{--}100 \mu\text{m}$ in thickness, were taken from the rock samples. Polishing powder and distilled water only were applied to the slides during preparation. Any contact with organic substances was carefully avoided.

Thirteen well preserved specimens were selected for the analysis. They were all located below the surface cut and were completely enveloped by the mineral matrix. Separated analyses were made of different portions of the individuals—the cell interior, the cell wall and the enveloping sheath. Control measurements were extended to the mineral vicinity of the specimens and to the surface of the slide.

Two different types of spectra were obtained. One (Fig. 4) is typical of the dark brownish substance composing sheath, cell

wall and cell filling (*s, p* in Fig. 1*a, e*), the other spectrum is typical (Fig. 5) of the cell vacuole (*v* in Fig. 1*c*).

Hydrocarbons of high aromatisation and condensation similar to those present in asphaltites are indicated by most of the spectra obtained from cell walls and fillings (Fig. 4*a-c*). Some spectra differ from this type by developing a stronger intensity at $1,360 \text{ cm}^{-1}$ (Fig. 4*e, f*). Such patterns are characteristic of

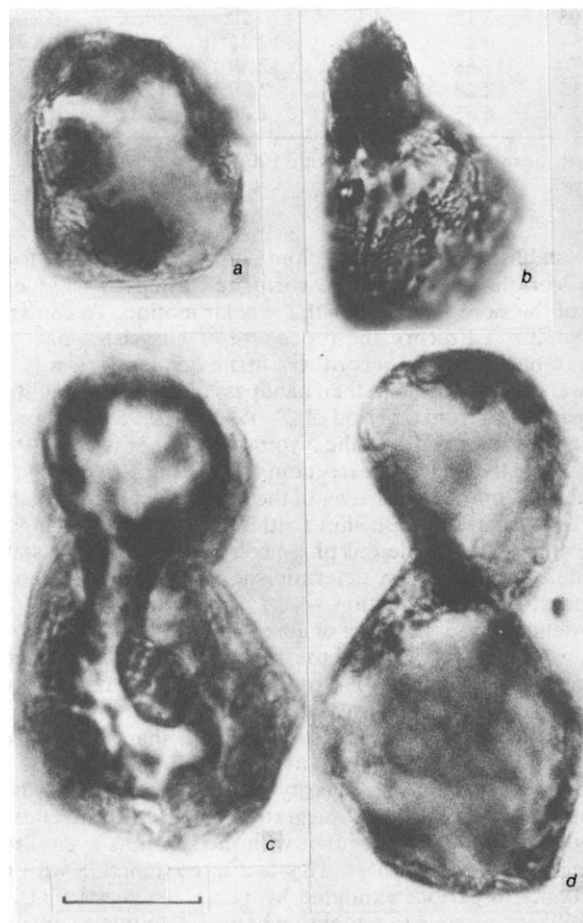


Fig. 2 *Isuasphaera* in different stages of budding (compare Fig. 3). Scale bar, 10 μm .

amorphous carbon produced by charring of organic material. It is therefore possible that some graphite may also be present in the *Isuasphaera* matter. The Raman spectrum of graphite⁵ is characterised by a strong, sharp line at $1,580 \text{ cm}^{-1}$, and clearly differs from the larger band formed in the *Isuasphaera* spectra at about $1,595 \text{ cm}^{-1}$. Another difference is the band at $1,360 \text{ cm}^{-1}$ which seems to be narrower in the graphite spectrum than in the *Isuasphaera* spectrum. Thus, graphite can account for, at most, only part of the fossil matter. It can be concluded from the analyses that *Isuasphaera* consists of organic material which is partly present in a carbonised condition, partly in a high rank of coalification very similar to a final stage of graphitisation. This is in accordance with the metamorphic condition of the enclosing rock which seems to be in the range of an upper greenschist facies⁶.

Characteristic Raman spectra were obtained from two cell vacuoles of *Isuasphaera* (Fig. 5). In these, we find the typical lines of organic functional groups. A stretching vibration of the carbonyl group $\text{C}=\text{O}$ appears at $1,745 \text{ cm}^{-1}$. The high wavenumber corresponds to that of an ester carbonyl. Indications of the presence of an ester are also found with the line at

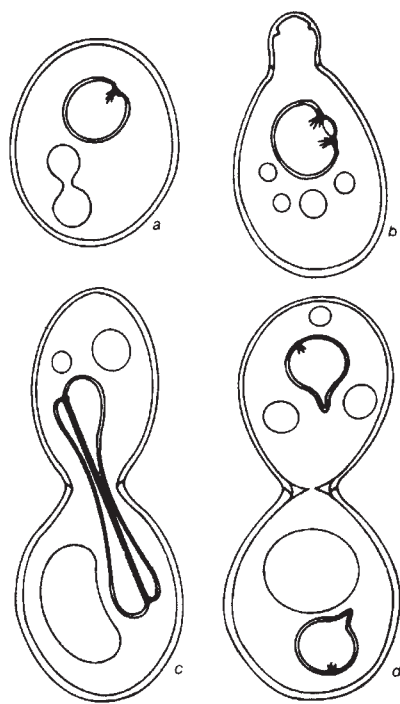


Fig. 3 Stages of a budding yeast cell: n, nucleus, v, gas vacuole (simplified after Matile *et al.*¹⁰).

850 cm^{-1} , which is the region of stretching vibration $\nu\text{C}=\text{C}$ adjacent to a carbonyl carbon bond. Another indication is the high frequencies of $\nu\text{C}-\text{H}$ vibrations near 3,000 cm^{-1} . An intense line at 1,460 cm^{-1} shows that a large amount of aliphatic hydrocarbons is present because the line corresponds to the deformation vibrations of CH_3 and CH_2 groups. The same groups are responsible for the broad signal at about 3,000 cm^{-1} which is a region in which the $\text{C}-\text{H}$ stretching vibrations appear. The highest peak is a doublet at 2,950 and 2,967 cm^{-1} ; a position which is typical for $\text{C}-\text{H}$ vibrations in an electro-negative environment corresponding to that of an ester function.

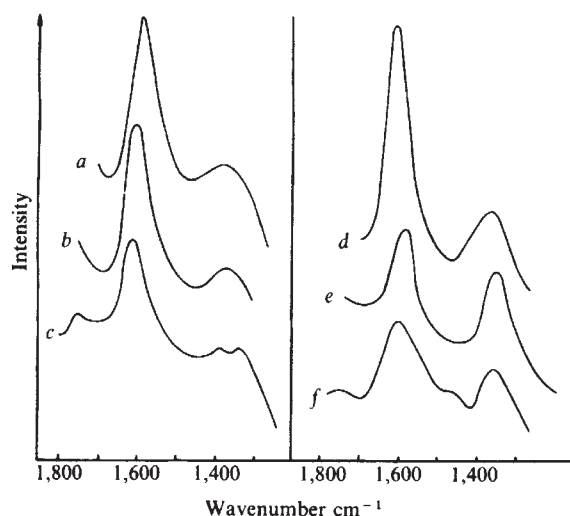


Fig. 4 Raman spectra obtained from cell wall and cell filling of *Isuasphaera* (b-f) compared with that of resinous asphaltite from Boscan (a).

Between 1,460 and 1,745 cm^{-1} , no significant lines for $\text{C}=\text{C}$ double bonds occur.

We believe that the preservation of the vacuoles is due to the special conditions in the cherts. Normally, the volatile compounds set free from the mother bitumen during metamorphism escape through pores and cracks of the rock. However, degassing cannot take place from material which is hermetically sealed in a quartz grain. Here the volatile fractions produced remain trapped in the matrix and after cooling, condense at the place of their origin, that is, at the cell lumen or nearby.

Structural comparisons have shown² that *Isuasphaera* conforms with recent asporogenous yeasts in all its preserved morphological properties, such as size, shape, structure of the cell wall and sheath, mode of growth and reproduction (Figs 2 and 3). But in the present context budding is not an exclusive

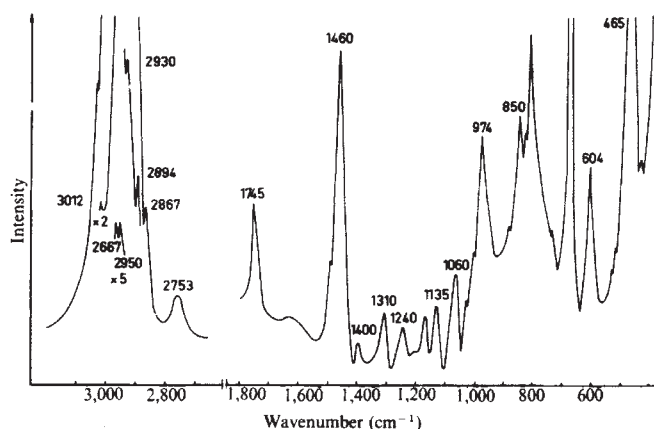


Fig. 5 Raman spectrum obtained from the cell vacuole of *Isuasphaera* (v in Fig. 1c).

characteristic of living cells. Multilateral budding, which is not dissimilar to that of *Isuasphaera* is, for example, known from Fox's microspheres—the artificial protein aggregates produced by certain laboratory experiments (Fig. 6). Because of their interesting budding abilities, the microspheres are often considered as a possible model for the hypothetical protocell⁸.

However, fossil *Isuasphaera* differs from the microspheres in a series of specific features showing that it was not an abiotic product but a true organism. The exterior multilaminate sheath enveloping the *Isuasphaera* cell can only be understood as a product of biological activity.

Metabolic production is also indicated by the esters filling the cell vacuole of *Isuasphaera* (Fig. 4). Similar fat globules are found in recent yeast cells. They commonly make up 30–40% of the dry weight of the organism and reach 50–60% in some genera. Compounds of the lipid fraction are known to be resistant to decomposition and can, in the proper environment, last for long periods of time. They are also known to be a main source of the petroleum hydrocarbons.

There is little doubt that *Isuasphaera* is an organism. This is indicated by the morphological and chemical details. A problem results from the similarities between *Isuasphaera* and recent yeasts. If we conclude that yeasts existed during Isua times, life must be much older than 3,800 Myr. According to our present knowledge, the Earth's sphere formed about 4,450 Myr ago⁹. Consequently biophile conditions, implying a solid crust, cool temperatures and liquid water cannot have been realised on Earth earlier than about 4,400 Myr ago. The time span of roughly half a billion years, however, appears too short for the evolution from a simple organic compound to a eukaryotic organism to have occurred.

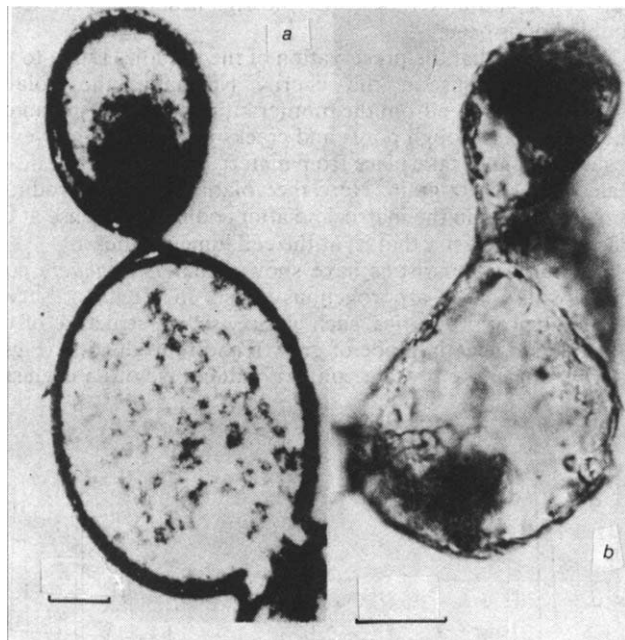


Fig. 6 *Isuasphaera* (b) compared with Fox's microsphere (a), an artificial budding structure consisting of protein, which has often been considered as a model of the hypothetical protocell (a from 8). Scale bars in a, 0.5 μm ; b, 10 μm .

However, the classification of *Isuasphaera* to yeasts or other eukaryotes is uncertain. The question of whether an organism is an eukaryote or not can hardly be decided without knowledge of the protoplasm which no longer shows structural details in its fossil condition. The example of the microspheres (Fig. 6) indicates that cellular budding does not have to be a very complicated process nor does it have to depend on a complicated endocellular apparatus.

Thus, it seems that *Isuasphaera* in spite of its yeast-like appearance may occupy an evolutionary level far below that of eukaryotes. Consequently, *Isuasphaera* may represent a half-way line between a microsphere-like protobiont and subsequent evolution.

We thank Dr P. W. U. Appel, and Dr M. Schidlowski, Mainz, FRG, for the rock samples. The work is part of the DFG-project Pf 21/27.

H. D. PFLUG

Geological Institute,
Justus Liebig University,
63 Giessen, FRG

H. JAECHKE-BOYER

Laboratoire d'Application Jobin Yvon,
91160 Longjumeau, France

Received 30 April; accepted 11 June 1979.

1. Moorbath, S., O'Nions, R. K., Pankhurst, R. J. *Earth planet. Sci. Lett.* **27**, 229–239 (1975).
2. Pflug, H. D. *Naturwissenschaften* **65**, 611–615 (1978).
3. Venec-Peyre, M. T. & Jaeschke-Boyer, H.: *C. r. hebdom. Séanc. Acad. Sci., Paris* **287**, 607–609 (1978).
4. Delhay, M., Dhamelincourt, P. J. *Raman Spectrosc.* **3**, 33–43, (1975).
5. Tuinstra, F. & Koenig, J. J. *Chem. Phys.* **53**, 1126 (1970).
6. Nagy, B., Zumberge, J. E. & Nagy, L. A. *Proc. natn. Acad. Sci. U.S.A.* **7**, 1206–1209 (1975).
7. Friedel, R. A. & Carlson, G. L. *Fuel* **51**, 194–198 (1972).
8. Fox, S. W. *Naturwissenschaften*, **56**, 1–9 (1969).
9. Schidlowski, M., Appel, P. W. U., Eichmann, R. & Junge, C. E. *Geochim. cosmochim. Acta* **43**, 189–199 (1979).
10. Matile, H., Moor, H. & Robinow, C. F. in *The Yeasts*, Vol. 1 (eds Rose, A. H. & Harrison, J. S.) 219–302 (Academic, New York, 1969).

An early Cretaceous terrestrial crocodilian and the opening of the South Atlantic

THE discovery in the lower Cretaceous of Niger of *Araripesuchus*, a small terrestrial crocodilian previously known only from Brazil, corroborates the existence of an African–South American faunal community until the Aptian. We show here that this provides additional evidence that an open seaway connecting the northern and southern parts of the Atlantic was not established until well into the early Cretaceous.

Among the abundant remains of terrestrial and freshwater vertebrates found (by P.T.) in the vast early Cretaceous fossil locality of Gadoufaoua (northeastern part of the Republic of Niger)¹ are those of a small short-snouted crocodilian. They were first discovered in 1966, and have been mentioned¹ as those of a 'brevirostrine crocodilian'. These specimens include incomplete skulls and lower jaws, vertebrae, dermal scutes, and limb bones. After acid preparation (by E.B.) of the very hard sandstone matrix, they were first found to belong to the genus *Araripesuchus*, a notosuchian crocodilian previously recorded by Price² from the early Cretaceous of Brazil.

The most significant specimen is a partial skull with the lower jaw still attached (Fig. 1); the part of the head located behind the orbits has disappeared. This skull is small: its complete length must have been about 12 cm, which indicates a crocodilian about 90 cm long. The snout is remarkably short and comparatively broad. The external nostrils are located in a very anterior position. The maxillae are somewhat 'swollen' behind their junction with the premaxillae. There is a large well defined antorbital fenestra, bordered anteriorly by the maxilla, posteriorly by the lacrymal. The orbits were large. A palpebral bone is present. In the upper jaw, there are 4 teeth in the premaxilla and 11 in the maxilla. The third premaxillary tooth and the third maxillary tooth are enlarged. The back teeth of the maxilla are small, laterally compressed, and have a rounded apex. The internal nares are large and located between the palatines and the pterygoids in a typically mesosuchian position. The lower jaw has a moderately long symphysis, in which the splenials take part. Postcranial remains indicate that the vertebrae are amphicoelous and that a dorsal and a ventral armour of dermal scutes were present.

These characters are consistent with the inclusion of this form in the mesosuchian infraorder Notosuchia and the family Uruguaysuchidae, as defined by Gasparini³. The Notosuchia are a group of small, short-snouted mesosuchian crocodilians, represented in South America by the families Uruguaysuchidae (early Cretaceous and beginning of the late Cretaceous) and Notosuchidae (late Cretaceous)³. The Libycosuchidae⁴, from the Cretaceous of Africa⁵, could also be referred to this infraorder. Closer comparison shows that the fossils from Gadoufaoua are very similar to the uruguaysuchid *Araripesuchus gomesii* Price, 1959, known from a skull and lower jaw found in the Santana Formation of northeastern Brazil². The resemblances (Fig. 2) are such that it seems safe to refer the African form to the genus *Araripesuchus*. However, some minor differences in the shape of the teeth (which in the African form are reminiscent of *Uruguaysuchus* from the Cenomanian of Uruguay⁶) and in their relative sizes, as well as in the shape of the antorbital fenestra, will probably justify the suggestion of a new species, closely related to *A. gomesii*. This is thus the first record of an uruguaysuchid in Africa, and it has significant palaeobiogeographical implications.

The Elrhaz Formation of Niger, which has yielded the African *Araripesuchus*, and the Santana Formation of Brazil, from which comes the only known specimen of *A. gomesii*, are approximately of the same age. According to stratigraphic and palaeontological evidence¹, the Elrhaz Formation is probably of late Aptian age. The age of the Santana Formation has been disputed, but recent research both on the invertebrate⁷ and verte-

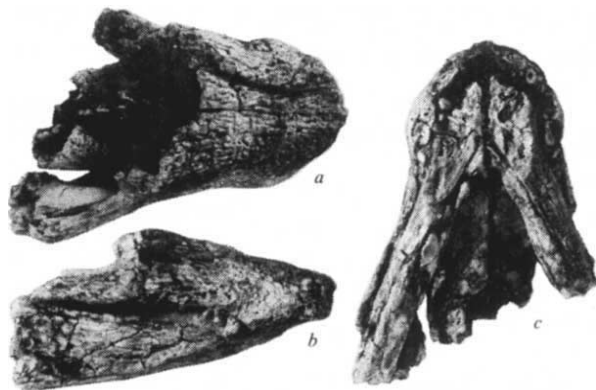


Fig. 1 *Araripesuchus* sp. from the Aptian Elrhaz Formation of Gadoufaoua, Republic of Niger. Incomplete skull and lower jaw collected by P.T. Museum National d'Histoire Naturelle, Paris, No. GDF 700. *a*, Dorsal view; *b*, lateral view, showing the antorbital fenestra; *c*, ventral view, showing the internal nares. Total length of specimen: 90 mm. Photographs by D. Serrette.

brate⁸ faunas tends to refer it to the late Aptian⁹. The discovery of *Araripesuchus* in the Aptian of Niger provides further evidence that the Santana Formation is not of late Cretaceous age, as was formerly believed.

The occurrence of very closely related species of *Araripesuchus* in the Aptian of Niger and Brazil also provides additional evidence of the existence of a terrestrial and freshwater vertebrate fauna little if at all divergent between Africa and South America until late in the early Cretaceous. This was already indicated by the presence of several very closely related forms (including freshwater fishes and the giant crocodilian *Sarcosuchus*) at Gadoufaoua and in the freshwater deposits, of probably Aptian age, of the Bahia/Reconcavo basin, on the eastern coast of Brazil¹⁰. We interpreted this fact as evidence that South America and Africa had not yet completely separated in the Aptian¹⁰—that there was no complete marine connection between the North Atlantic and the South Atlantic at the time—a hypothesis supported by studies on the ostracod faunas of the early Cretaceous coastal basins of Gabon and Bahia¹¹. Note that the anatomy of the *Notosuchia* indicates that these

small crocodilians were adapted to a terrestrial, somewhat mammal-like, way of life; *Araripesuchus* certainly could not cross any large area of saltwater.

Evidence relating to the date of the opening of a continuous seaway between Africa and South America has been based largely on the distribution of ammonites. On the basis of these kind of data, Reymont¹² concluded that a complete marine connection allowing a faunal interchange became established in the middle Turonian. Kennedy and Cooper¹³ have then claimed that such a connection existed as early as the late Albian, a conclusion also supported by Forster¹⁴ and Forster and Scholz¹⁵. Palaeobathymetric¹⁶ and sedimentological¹⁷ studies support the hypothesis of the establishment of a marine connection at the end of the early Cretaceous. These views, which suggest that a seaway separating Africa from South America came into being in the Albian, about 110 Myr ago, are consistent with our conclusions based on the Aptian terrestrial faunas of these continents. However, in a recent collection of palaeocontinental maps¹⁸, Smith and Briden favour an earlier date—the beginning of the Cretaceous—for the opening of the South Atlantic and the separation of Africa from South America. We believe that palaeontological evidence based on the distribution of terrestrial and freshwater vertebrates, reinforced by the discovery of the crocodilian *Araripesuchus* in Niger, shows that a land connection between South America and Africa may still have existed as late as Aptian times. This obviously favours a comparatively late opening of the South Atlantic, at the end of the early Cretaceous.

ERIC BUFFETAUT

Laboratoire de Paléontologie des Vertébrés,
Université Paris VI,
4 place Jussieu,
75230 Paris Cedex 05, France

PHILIPPE TAQUET

Institut de Paléontologie,
8 rue de Buffon,
75005 Paris, France

Received 11 May; accepted 22 June 1979.

1. Taquet, P. *Géologie et Paléontologie du Gisement de Gadoufaoua (Aptien du Niger)* (Cahiers de Paléontologie, Editions du CNRS, Paris, 1976).
2. Price, L. I. *Bolm. Div. Geol. Miner., Bras.* **188**, 1–55 (1959).
3. Gasparini, Z. B. de, *Ameghiniana* **8**, 83–103 (1971).
4. Stromer, E. *Abh. Bayer. Akad. Wiss., Math.-phys. Kl.* **27**, 1–16 (1914).
5. Buffetaut, E. *Mitt. Bayer. Staatssamm. Paläont. hist. Geol.* **16**, 17–28 (1976).
6. Rusconi, C. *Bol. Inst. Geol. Montev.* **19**, 1–64 (1933).
7. Beurlen, K. *Anais Acad. bras. Cienc.* **38**, 455–464 (1966).
8. Silva Santos, R. da & Gomes Valença, J. *Anais Acad. bras. Cienc.* **40**, 339–360 (1968).
9. Bonaparte, J. F. *El Mesozoico de America del Sur y sus Tetrapodos* (Opera Lilloana 26, San Miguel de Tucuman, 1978).
10. Buffetaut, E. & Taquet, P. *Palaeontology* **20**, 203–208 (1977).
11. Krommelbein, K. *Proc. 2nd Gondwana Symp. South Africa 1970*, 617–619 (Pretoria, 1971).
12. Reymont, R. A. *Rev. Géogr. phys. Géol. dyn.* **16**, 61–70 (1974).
13. Kennedy, W. J. & Cooper, M. J. *Geol. Soc. Lond.* **131**, 283–288 (1975).
14. Forster, R. *Nature* **272**, 158–159 (1978).
15. Forster, R. & Scholz, G. *Neues Jb. Geol. Paläont. Mh.* **157**, 109–119 (1979).
16. Sclater, J. G., Hellinger, S. & Tapscott, C. J. *Geol.* **85**, 509–552 (1977).
17. Van Andel, T. H., Thiede, J., Sclater, J. G. & Hay, W. W. J. *Geol.* **85**, 651–698 (1977).
18. Smith, A. G. & Briden, J. C. *Mesozoic and Cenozoic Paleogeographic Maps* (Cambridge University Press, 1977).

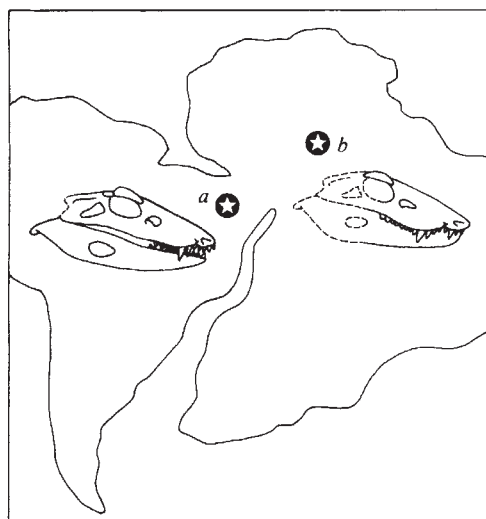
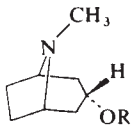
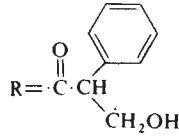
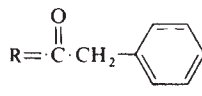


Fig. 2 Geographical distribution of the genus *Araripesuchus* in the Aptian. *a*, *Araripesuchus gomesii* Price, 1959, from the Santana Formation, Chapada de Araripe, northeastern Brazil (after Price²). *b*, *Araripesuchus* sp., from the Elrhaz Formation of Gadoufaoua, Tégama basin, Niger. Length of the lower jaw in both about 130 mm. Position of South America and Africa in the late Aptian after Forster¹⁴.

Toxicity induced in the tobacco horn-worm (*Manduca sexta* L.) (Sphingidae, Lepidoptera)

INSECTS feeding on toxic plants demonstrate three fairly distinct types of life-style^{1–3}. There are cryptic species, which metabolise or rapidly excrete the toxic substances present or avoid their ingestion by selective feeding; aposematic or warningly coloured species which store plant toxins in their tissues unchanged or slightly modified⁴; and aposematic species which superficially resemble or mimic toxic species—without actually storing poisonous plant products—or those warningly-coloured

Table 1 Alkaloids from insect and plant

Compound	Pupae	Larval frass	<i>Atropa belladonna</i> L.
			
I	+++	+++	+++
	-	++	-
II	-	++	-
	-	(Trace)	-
III	-	(Trace)	-

Symbols represent relative amounts determined by gas liquid chromatography¹⁰.

non-storers which secrete their own toxins. It is generally agreed that the cryptic life-style is more 'successful' than either of the other two, which are relatively rare, and it is difficult to envisage the evolutionary steps necessary to enable a species to change to the more hazardous warning life-style. If, however, circumstances favour a switch to certain toxic host plants, a cryptic insect is frequently destined to become warningly coloured^{1,5}. On the basis of experiments with *Manduca sexta*, the tobacco horn-worm, we suggest that the evolution of an aposematic poisonous insect, from the more common, harmless, cryptic type, may simply involve a change to a related food plant containing different toxic properties from those of its usual host. Storage and the acquisition of toxicity and warning colour could follow this crucial switch.

In the Sphingidae there are several examples of aposematic larvae which feed on toxic plants, such as *Plumeria* (ref. 6 and unpublished results of M.R. and N.M.), but which give rise to cryptic, non-poisonous adults. This suggested that storage, even if temporary, sometimes occurs in this group. It seemed more likely that a cryptic insect such as *Manduca*, which excretes toxins rapidly^{7,8}, rather than those which metabolise them, would be able to switch more readily to tissue storage.

We reared *M. sexta* (from hatching to the pupal stage) on the abnormal food plant belladonna (*Atropa belladonna* L.), and controls on leaves of tobacco (*Nicotiana tabacum* L.) or standard Yamamoto⁹ artificial diet lacking leaf powder. Between one and three of the pupae resulting from belladonna-reared larvae were then fed to each of six hens, all of which died within 1–7 d. Two of the tobacco-reared specimens, and 14 of those fed on artificial diet (controls) were eaten by two hens, neither of which suffered any visible ill effects and continued egg-laying throughout the 7 d of the experiment.

Extracts of four belladonna-reared pupae (3.354–4.207 g each) were injected intraperitoneally into laboratory mice, of which three out of four died within 20–25 h⁶. Extracts of pupae reared on tobacco and artificial diet failed to elicit any visible response in the mice.

Analysis of the belladonna-reared pupal stage revealed the presence of atropine in the body tissues (Table 1) and both atropine (I) and two metabolites (II and III) were found in larval frass. The pupae of *M. sexta*, when reared on belladonna, therefore contain lethal toxins.

We thank Carroll Williams for *M. sexta* eggs and Brian Gardiner and Catherine Godfrey for assistance with rearing caterpillars.

MIRIAM ROTHSCILD

Ashton Wold,
Peterborough, UK

ROBIN APLIN
JOHN BAKER

Dyson Perrins Laboratory,
South Parks Road,
Oxford, UK

NEVILLE MARSH

Department of Physiology,
Queen Elizabeth College,
Campden Hill Road,
London W8, UK

Received 2 April; accepted 5 July 1979.

1. Rothschild, M. in *Insect Plant Relationships* (ed. van Emden, H. F.) 59–83 (Blackwell, Oxford, 1972).
2. Rothschild, M. in *Phytochemical Ecology* (ed. Harborne, J. B.) 2–12 (Academic, New York & London, 1972).
3. Rothschild, M. & Reichstein, T. *Nova Acta Leopoldina*, Suppl. 7, 1–31 (1975).
4. Dixon, C. A., Erickson, J. M., Kellett, D. N. & Rothschild, M. *J. Zool., Lond.* **185**, 437–467 (1978).
5. Rothschild, M. in *Ecological Genetics and Evolution* (ed. Creed, G. R.) 202–223 (Blackwell, Oxford, 1971).
6. Marsh, N. & Rothschild, M. *J. Zool., Lond.* **174**, 89–122 (1974).
7. Hodgson, E., Self, L. S. & Guthrie, F. E. *Proc. twelfth int. Cong. Ent.* 210 (1975).
8. Yang, R. S. H. & Guthrie, F. E. *Ann. ent. Soc. Am.* **62**, 141–146 (1969).
9. Yamamoto, R. T. *J. econ. Ent.* **62**, 1427–1431 (1969).
10. Paik, N. H., Park, M. K. & Kim, M. S. *Soul Taehakkyo Yakhak Nonmunjip* **1**, 26–39 (1976); (*Chem. Abs.*, **87**, 90771a, 1977).

Retinal site of transient tritanopia

WHEN the eye is adapted to a bright yellow light, a marked loss of sensitivity for short-wavelength stimuli occurs in the first few seconds after extinction of this light. This phenomenon was first reported by Stiles¹ and has been termed transient tritanopia by Mollon and Polden². The latter authors suggest that this apparent suppression of the short-wave sensitive cones (hereafter called blue cones) during rapid dark adaptation is effected through inhibition by the recovering long- and middle-wave sensitive cone mechanisms² (hereafter called red and green cones, respectively). In this study we attempted to localise the site of blue cone inhibition through intra-retinal recordings from the intact rhesus monkey eye. We found transient tritanopia to be absent at the receptor level but present at the b-wave level, and we conclude that transient tritanopia originates at a site between the receptors and the bipolar cells, probably at the horizontal cells, as was earlier hypothesised by Augenstein and Pugh³.

The signal that is recorded with a microelectrode positioned in the central fovea of a rhesus monkey eye, the local electroretinogram (LERG), consists mainly of the positive-going cone late receptor potential (LRP), also termed PIII, with a superimposed negative-going b-wave⁴, representing activity at the bipolar cell level^{5,6}. Figure 1 shows typical waveforms of responses to 447-nm and 551-nm flashes in two conditions. In the first condition, 'during yellow', a 10⁵-troland yellow/orange background was present. The second condition, 'after yellow' (the transient tritanopia condition), is in the dark, shortly after extinction of the yellow/orange background. In the lower panels (Fig. 1), the 551-nm flash response 'after yellow' shows a large increase of both receptor potential and b-wave, compared with the 'during yellow' condition (note the difference in vertical scale). This increase reflects the removal of a desensitising field, that is, switching off the background. The situation is different for the 447-nm flash responses (Fig. 1, upper panels), where the removal of the background field leaves the receptor potential unaltered, but virtually abolishes the b-wave.

To assess which cone systems are involved in evoking these responses, the spectral sensitivities of both LERG components in the two conditions were determined for three monkeys (Fig. 2a, b). Spectral sensitivity is defined as the inverse of the flash intensity necessary to evoke a criterion response. Figure 2a gives the receptor potential data. The data points were fitted with the blue cone spectral sensitivity function⁷ in the short-wavelength region and the CIE standard photopic spectral sensitivity function for a 10° stimulus⁸ ($V_{10,\lambda}$) in the rest of the spectrum. The low sensitivity of the long- and middle-wavelength-sensitive cones relative to the blue cones is apparent

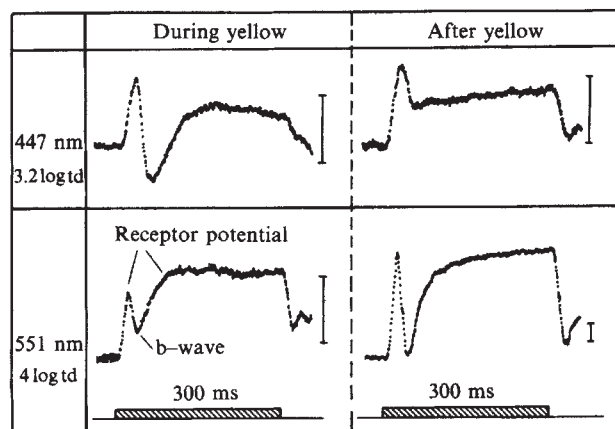


Fig. 1 Responses obtained from the central fovea of the intact rhesus monkey retina with a technique adapted from Baron and Boynton¹⁵. A microelectrode, protected by a guide needle, was driven through the sclera and directed to the macular region under direct observation. The tip of the electrode (~1 μm) was positioned in the outer segment layer of the central fovea and a reference electrode was placed in the vitreous at 1 mm from the retina. The responses are averages of 90 flashes, each lasting 300 ms; repetition time was 1,200 ms. Vertical calibration bars represent 50 μV. The left-hand panels show responses to 447-nm and 551-nm flashes presented on a 10° troland yellow/orange background (Schott OG 570 cut-off filter). Stimulus intensities were 3.2 log td and 4 log td, respectively. Both waveforms have a distinct b-wave. The right-hand panels show responses to exactly the same stimuli, 1 s after extinction of the background. At 447 nm, top right, the b-wave has disappeared, whereas at 551 nm, bottom right, the whole response has increased in amplitude (note the smaller calibration bar). For the purpose of response averaging in the 'after yellow' condition, the background was repeatedly turned on and off, the on-phase lasting 10 s, the off-phase 1.6 s. In each dark interval one 300-ms flash was given 1 s after extinction of the background. This cycle was repeated until 90 responses were averaged. The extinction of a 10° troland yellow background field produces a large off-response that contaminates the recordings in the 'after yellow' condition. Therefore, this off-response was recorded separately, averaged, stored and subtracted from each averaged response in the 'after yellow' condition. Both stimulus and background field subtended 10°, centred around the fovea. The receptive field of our microelectrode is about 8°. Responses from stray light beyond the 10° stimulus field therefore do not contribute to the signal.

in the 'during yellow' condition and this is usually called chromatic adaptation. Turning off the yellow/orange background leaves the blue cone sensitivity unaltered (see the 'after yellow' data points), but the long-wave sensitivity increases by about 1.6 log units which is as expected after removal of a yellow/orange chromatic adapting field. Figure 2b shows the b-wave spectral sensitivity data. The 'during yellow' data at short wavelengths are again fitted with the macaque blue cone spectral sensitivity function⁷, but the long- and middle-wavelength points are now fitted with the green cone spectral sensitivity function⁹ (this will be discussed below). In the 'after yellow' condition the sensitivity at short wavelengths is about 1.5 log units lower than in the 'during yellow' condition, which is indeed transient tritanopia: shortly after extinction of a yellow/orange background there is a marked decrease in sensitivity for blue stimuli¹⁰. The middle- and long-wave sensitivity increases by about 1 log unit when the background is removed.

To determine which cone mechanisms mediate the b-wave spectral sensitivity 'after yellow', we tried to fit the $V_{10,\lambda}$ function and the blue cone spectral sensitivity function to the data. However, a good fit could not be obtained. Apparently, the blue cone system does not contribute to the response. The green cone spectral sensitivity function⁹, however, gives a very satisfactory fit between 400 and 620 nm. The sensitivity at 660 nm is clearly higher than predicted by the green cone sensitivity curve but this is probably caused by a contribution of the red cone system to the response. These findings corroborate the psychophysical results of Mollon and Polden², who describe the short-wavelength spectral sensitivity (<500 nm) in the 'after yellow' condition using Stiles' π_4 (green sensitive) mechanism. It is indeed surprising that the green cone spectral sensitivity function, which certainly does not fit the receptor potential data, fits the b-wave data points 'during yellow' and 'after yellow' in the above mentioned spectral regions, and that the $V_{10,\lambda}$ function does not. This might be explained by previously reported colour-opponent features of the b-wave^{11,12}.

Pugh and Mollon¹³ recently proposed a two-site adaptation model for Stiles' π_1 (blue sensitive) mechanism that includes an explanation of transient tritanopia. The first site is in the receptors themselves and the second site is believed to be in the retina proximal to the receptors. This second site is thought to have blue-yellow colour-opponent properties. It is supposed that sensitivity at the second site depends on the balance of the blue and yellow signals, being least when the site is polarised. To account for transient tritanopia it is suggested that a restoring force, which acts to reduce the polarisation of the site, is built up slowly during adaptation to the yellow field and continues to act for some time after extinction of the field, so producing a polarisation in the opposite sense. The present results fit this view rather well: the operational inhibition of the blue cone signal does not take place at the receptor level, but at some more proximal site. As the suppression is manifest at the level of the b-wave, it must take place at the bipolar cells⁶ or earlier. The horizontal cells, which are distal to the bipolar cells, are the most likely candidates for blue cone inhibition: first, because of their anatomical suitability (they make cross-connections between various receptor cells) and second, because, at least in lower vertebrates, they are chromatically coded¹⁴.

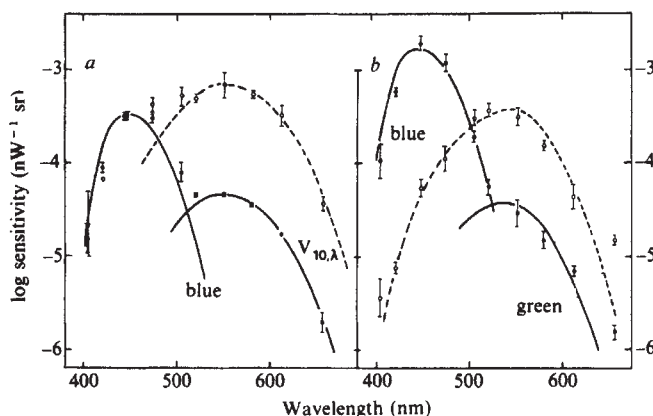


Fig. 2 a, Spectral sensitivity of the receptor activity, measured on a yellow/orange background (●) and 1 s after extinction of this background (○). Data are averages of three experiments. Receptor activity, the LRP or proximal PIII, was estimated by measuring the slope of the leading edge of the LERG response. Taking the initial slope allows consideration mainly of distal PIII without contamination of subsequent proximal PIII, onset of the b-wave potential or feedback from higher order cells¹⁶. Criterion for sensitivity was a response of 2 μV ms⁻¹. Error bars represent ±1 s.e.m. b, Spectral sensitivity of the b-wave 'during yellow' (●) and 'after yellow' (○). The b-wave amplitude was defined as the distance between initial positive peak and first negative deflection. Criterion was 45 μV and error bars represent ±1 s.e.m. This figure shows transient tritanopia: the sensitivity for blue (short wave) stimuli is much reduced when the yellow/orange background is removed ('after yellow' data points), compared with when the background is present. This is in contrast to a, where the blue sensitivity 'during yellow' and 'after yellow' is the same.

In conclusion, our electrophysiological data confirm the hypotheses based on psychophysical studies^{2,3} in two ways: short-wavelength flashes, after switching off a yellow adaptation light, are 'seen' by the green cones rather than by the blue cones; and the inhibition of the blue cones can probably be located at the horizontal cell level.

This study was supported by the Netherlands Organization for the Advancement of Pure Research.

J. MATHEE VALETON
DIRK VAN NORREN

*Institute for Perception TNO,
Kampweg 5, 3769 DE Soesterberg,
The Netherlands*

Received 8 May; accepted 20 June 1979.

1. Stiles, W. S. *Documenta Ophthalm.* **3**, 138–163 (1949).
2. Mollon, J. D. & Polden, P. G. *Nature* **258**, 421–422 (1975); **259**, 570–572 (1976); *Phil. Trans. R. Soc.* **278**, 207–240 (1977); *Vision Res.* **19**, 435–440 (1979).
3. Augenstein, E. J. & Pugh, E. N. *J. Physiol., Lond.* **272**, 247–281 (1977).
4. Brown, K. T. *Vision Res.* **8**, 633–677 (1968).
5. Miller, R. F. & Dowling, J. E. *J. Neurophysiol.* **33**, 323–341 (1970).
6. Arden, G. B. & Brown, K. T. *J. Physiol., Lond.* **176**, 429–461 (1965).
7. Norren, D. van & Padmos, P. *Vision Res.* **13**, 1241–1254 (1973).
8. Wyszecki, G. & Stiles, W. S. *Color Science* (Wiley, New York, 1967).
9. Vos, J. J. & Walraven, P. L. *Vision Res.* **11**, 799–818 (1971).
10. Valeton, J. M. & Norren, D. van *Vision Res.* (in the press).
11. Padmos, P. & Graf, V. *Proc. XIth ISERG Symp. Doc. Ophth., Proc. Ser.*, 307–314 (Junk, The Hague, 1974).
12. Norren, D. van & Baron, W. S. *Vision Res.* **17**, 807–810 (1977).
13. Pugh, E. N. & Mollon, J. D. *Vision Res.* **19**, 293–312 (1979).
14. Naka, K. I. & Rushton, W. A. H. *J. Physiol., Lond.* **185**, 536–555 (1966a).
15. Baron, W. S. & Boynton, R. M. *Vision Res.* **14**, 495–501 (1974).
16. Fulton, A. B. & Rushton, W. A. H. *Vision Res.* **18**, 785–792 (1978).

Rapid effect of change in cerebrospinal fluid sodium concentration on salt appetite

THIRST can be evoked within 1–2 min by intracarotid infusion of hypertonic NaCl. In contrast, the onset of a specific appetite for sodium salts as a result of rapid loss of body sodium is slow. A parotid fistula in naive sheep can produce a significant Na deficit (200–700 mmol) within 1–3 d but a Na appetite usually develops after 2–5 d (ref. 1). With large urinary water and Na loss induced by frusemide, water intake occurs within an hour but Na appetite is delayed until 24 h (ref. 2). Moreover, in the Na-depleted sheep or goats, increase in the Na concentration of cerebral arterial blood by 10–30 mmol l⁻¹ by slow intracarotid infusion of 4 M NaCl for 7 min before or during presentation of Na solution does not reduce Na intake^{3,4}. However, rapid intravenous infusion of significant amounts of isotonic and 2–4 M NaCl solutions in Na-deficient sheep substantially reduces Na appetite after 40–120 min^{5,6}. In the rat, tissue fluid sequestration produced by subcutaneous polyethylene glycol injection causes water drinking within an hour, but onset of salt appetite is delayed by 6–10 h.⁷ This delay also occurs in salt appetite produced by peritoneal dialysis⁸, tourniquet release⁹ and formalin injection^{10,11}. New data indicate the powerful effects of the hormones involved in the reproductive process in inducing salt appetite. With pregnancy and lactation in Na-replete rabbits, avid salt appetite develops with a daily turnover of 50% or more of extracellular fluid (ECF) Na (ref. 12). This is caused by the combined action of oestradiol, cortisol, ACTH, prolactin and oxytocin¹³. ACTH seems to act directly on the brain¹⁴. Recent experimental data provide new information on the central nervous system control mechanism. Weisinger *et al.*¹⁵ reported that in Na-depleted sheep, the systemic influx of hypertonic NaCl significantly reduced motivation for salt within 10–20 min. In the light of these data, we decided to evaluate the effect on Na appetite of increasing the cerebrospinal fluid (CSF) Na concentration or osmolality in Na-deficient sheep. We now report the first experiments to show that it is feasible to

induce salt appetite within an hour or less by a physiological procedure.

Sheep with a parotid fistula were trained to bar press for 600 mM NaHCO₃ (15 ml per delivery = 9 mmol Na). They were allowed to bar press for 2 h daily (1200–1400 h), and thus had been deprived of Na for 22 h (deficit of 350–450 mmol) when the bar press was made available. On the day of the experiment, the infusion (1 ml h⁻¹) into a lateral brain ventricle was begun 1 h before and continued until the end of the 2-h bar press period. The composition of artificial CSF has been described previously^{16,17}. CSF was sampled immediately after the end of infusion and analysed for [Na], [K] and osmolality. Data were analysed by a two-way analysis of variance with subsequent *t*-tests.

The results (Fig. 1, *n* = 6) show that within 10 min, intraventricular infusion of CSF [Na, 500 mM] caused a significant decrease in Na intake relative to infusion of CSF [Na, 150 mM]. When an infusion of 0.7 M mannitol–CSF [Na, 150 mM] was used to test the effect of increasing CSF osmolality without increasing CSF [Na], Na intake was almost doubled (Fig. 1). The increase relative to the CSF [Na, 150 mM] infusion was significant within 60 min and thereafter. In experiments where no i.v.t. infusion was given, the intake was similar to that with CSF [Na, 150 mM]. In five animals tested Na appetite was significantly increased within 60 min whether the mannitol infusion was begun 60, 15 or 0 min before the bar press period. No significant increase in water intake occurred. The effect of mannitol infusion on Na intake was the same whether water was concomitantly available to the animal or not.

Relative to infusion of CSF [Na, 150 mM], CSF (Na, 500 mM) increased CSF [Na] and therefore osmolality and decreased CSF [K]. Infusion of 0.7 M mannitol–CSF [Na, 150 mM] decreased CSF [Na] and [K], and increased CSF osmolality (Table 1).

The question of whether hypertonic mannitol would induce a Na appetite in Na-replete sheep was then examined. Naive Na-replete sheep with no known history of Na deficiency were given continuous access to water. Two-hour access to a bin containing 600 mM NaHCO₃ was given daily for a control period of 2–4 weeks. After two or three control infusions of artificial CSF [Na, 150 mM], which had no significant influence on NaHCO₃ intake (Fig. 2), the animals (*n* = 5) were infused on three consecutive occasions with 0.7 M mannitol–CSF [Na, 150 mM] (Fig. 2). This was usually done on three consecutive weeks. A 3-h infusion (1 ml h⁻¹) into a lateral brain ventricle was begun 1 h before access to Na.

Na intake increased in all animals on each of the three mannitol–CSF infusion days relative to the mean Na intake on the two days preceding the infusion (*P* < 0.01) in all experiments (Fig. 2). The following day intake was basal or below. A small increase in water intake occurred during the first hour of infusion on trial 3 (*P* < 0.02) but was never increased during the 2-h period when NaHCO₃ was available. In 18 other experiments with a cafeteria of Na, K and Mg salts the specificity of choice of Na salts was clearly shown (300–500 mmol increase,

Table 1 Effect of a 3-h i.v.t. infusion of various solutions on CSF composition

Condition	Na (mM)	K (mM)	Osmolality (Mosmol)
CSF [Na, 150 mM]	151.2 ± 0.4 (25)	2.84 ± 0.03 (25)	297 ± 1.3 (15)
0.7 M Mannitol CSF [Na, 150 mM]	129.9 ± 1.4‡ (28)	2.52 ± 0.02‡ (28)	379 ± 7.8‡ (17)
0.3 M Mannitol CSF [Na, 0 mM]	127.0 ± 2.8‡ (5)	2.72 ± 0.11 (5)	319 ± 7.2‡ (5)
CSF [Na, 500 mM]	177.2 ± 13.7‡ (4)	2.47 ± 0.08‡ (4)	

CSF was sampled from the ipsilateral ventricle immediately after the end of infusion. Statistical comparisons with CSF [Na, 150 mM] value by Student *t*-test (**P* < 0.05, ‡*P* < 0.01; ‡*P* < 0.001). Numbers in parentheses indicate the number of animals tested.

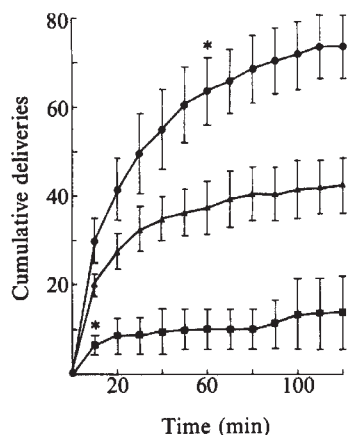


Fig. 1 Effect of i.v.t. (lateral ventricular) infusions on bar pressing for NaHCO_3 solution. The sheep were Na depleted as a result of parotid salivary loss for 22 h. Cumulative mean ($\pm$ s.e.m.) number of deliveries (15 ml 0.6 M $\text{NaHCO}_3 = 9$ mmol) over the 120-min period in the various conditions: $\blacksquare$, 3 h i.v.t. infusion (1 ml hr^{-1}) of artificial CSF [Na, 500 mM]; $\blacktriangle$, artificial CSF [Na, 150 mM]; $\bullet$, 0.7 M mannitol-artificial CSF [Na, 150 mM]. Infusion was started 1 h before the bar press period. $n = 6$ Na-deficient sheep. *Significantly different from CSF [Na, 150 mM] values, $P < 0.05$.

$P < 0.01$), although a small increase in KHCO_3 sometimes occurred.

With i.v.t. hypertonic mannitol the decrease in CSF [Na] and [K] was presumably due primarily to the osmotic withdrawal of water from the choroid plexus and surrounding brain tissue capillaries. A decrease of CSF [Na] with less water movement was obtained by infusion of 0.3 M mannitol-CSF [Na, 0 mM] ($n = 7$) (Table 1). The Na appetite induced was similar to that with 0.7 M mannitol-CSF [Na, 150 mM] (Fig. 2). No increase in water intake occurred.

The results of these experiments indicate that a change in the composition of CSF involving an increase or decrease in [Na] may rapidly alter salt appetite drive. A rapid induction, within 60 min or less, of a large Na appetite has been shown.

The induction of a Na appetite in the naive Na-replete sheep, augmentation in the moderately deplete animal, the demonstration by both voluntary intake and operant behaviour and the inhibition of appetite as a result of increase of CSF [Na] in the Na-deficient animal strengthen the inference that altered CSF [Na] affected the physiological control system. In all cases, the

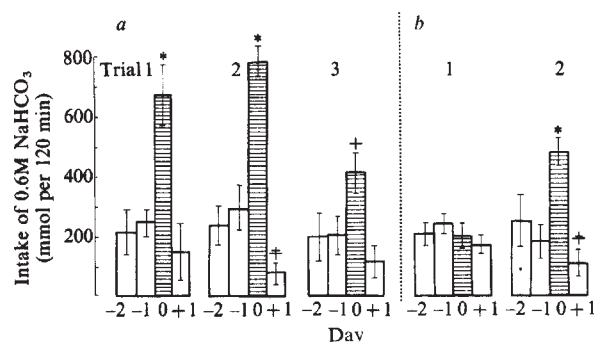


Fig. 2 *a*, Effect of 3 h i.v.t. infusion (1 ml h^{-1}) of 0.7 M mannitol-artificial CSF [Na, 150 mM] on voluntary intake of 0.6 M NaHCO_3 by 5 Na-replete sheep. Infusion was started 1 h before the 2 h access period on day 0. No infusion on days -2, -1 or +1. Water was continuously available. The effect of infusion on Na intake on three trials on consecutive weeks is shown. *b*, Effect in the same conditions of infusion of artificial CSF [Na, 150 mM] (1), 23 experiments in 13 sheep, and 0.3 M mannitol-artificial CSF [Na, 0] (2), 7 sheep. * $P < 0.001$; $\dagger < 0.01$; $\ddagger P < 0.02$.

experimentally induced changes in CSF [Na] were in the physiological range. The absence of significant water drinking despite an increased osmotic pressure may be due to concurrent decrease of CSF [Na]^{17,18}. With i.v.t. CSF [Na, 500 mM] infusions, if water was available during the whole infusion, or in the period 0–120 min when NaHCO_3 was available, drinking of 1–2 l occurred.

The results are consonant with the suggestion by D.A.D. that reduction in intracellular [Na] of the hypothalamic nerve cells subserving Na appetite is the basic mechanism for the genesis of this specific drive¹⁹; however, further experiments are required to confirm this. Whether the time interval for CSF changes to induce behaviour is too short for a transcription mechanism is also a matter for further investigation. Intracellular ionic change could act at the level of translation²⁰. A change in brain ECF [Na] may also have a direct effect on the membrane. The fact that the large salt appetite in the Na-replete animal during reproduction is generated by a sequence of steroid and peptide hormone actions²¹, presumably by the same neural mechanisms which respond to body Na deficit, is consistent with the idea that a complex neurochemical mechanism underlies the genesis of salt appetite.

R. S. WEISINGER
P. CONSIDINE
D. A. DENTON
M. J. MCKINLEY

Howard Florey Institute of Experimental
Physiology and Medicine, Melbourne,
Australia

D. MOUV

Department of Physiology,
University of Michigan, Ann Arbor,
Michigan 48109

Received 2 May; accepted 27 June 1979.

1. Denton, D. A. *Handbook of Physiology*, Section 6 Vol. 1 (ed. Code, C. F.) Ch. 31 (American Physiological Society, Washington, 1967).
2. Zimmerman, M. B., Stricker, E. M. & Blaine, E. H. *J. comp. Physiol. Psychol.* **92**, 501 (1978).
3. Beilharz, S., Bott, E., Denton, D. A. & Sabine, J. R. *J. Physiol., Lond.* **178**, 80–91 (1965).
4. Baldwin, B. A. *Physiol. Behav.* **16**, 59 (1976).
5. Beilharz, S. & Kay, R. N. B. *J. Physiol., Lond.* **165**, 468–483 (1963).
6. Denton, D. A., Orchard, E. & Weller, S. *Commun. Behav. Biol.* **6**, 245–258 (1971).
7. Stricker, E. M. & Wolf, G. *J. comp. Physiol. Psychol.* **62**, 275 (1966).
8. Ferreyra, M. & Chiaravaglio, E. *Physiol. Behav.* **19**, 197 (1977).
9. Stricker, E. M. in *Biological and Behavioural Aspects of Salt Intake* (eds Kare, M., Bernard, R. & Fregly, M.) (Academic, New York, in the press).
10. Stricker, E. M. & Wolf, G. *Ann. N.Y. Acad. Sci.* **157**, 553 (1969).
11. McCutcheon, B. & Levy, C. *Physiol. Behav.* **19**, 197 (1977).
12. Denton, D. A. & Nelson, J. F. *Endocrinology* **88**, 31–40 (1971).
13. Denton, D. A., McKinley, M. J., Nelson, J. F. & Weisinger, R. S. *Acta endocr.* **85**, S212, 3–4 (1977).
14. Blaine, E. H., Covelli, M. D., Denton, D. A., Nelson, J. F. & Shulkes, A. A. *Endocrinology* **97**, 793 (1975).
15. Weisinger, R. S., Denton, D. A. & McKinley, M. J. *J. comp. Physiol. Psychol.* **92**, 522 (1978).
16. Abraham, S. F. *et al.* *Acta Endocr.* **81**, 120–132 (1976).
17. McKinley, M. J., Denton, D. A. & Weisinger, R. S. *Brain Res.* **141**, 89–103 (1978).
18. Andersson, B. & Olsson, K. *Proc. 4th Int. Congr. Endocr.* **724**, (Excerpta Medica, 1973).
19. Denton, D. A. *Conditional Reflex* **1**, 144–170 (1966).
20. Douzou, P. & Maurel, P. *Trans. biochem. Sci.* **2**, 14 (1977).
21. Denton, D. A. & Nelson, J. F. *Endocrinology* **103**, 1880–1887 (1978).

Immunisation of guinea pigs and cattle against ticks

IXODID TICKS attach to the skin of their host and remain there for several days, engorging on blood and intermittently passing salivary secretions into the host¹. Several species of Ixodid ticks transmit pathogenic microorganisms along with their saliva into the host, and tick-borne diseases are of great economic importance, particularly in cattle in subtropical and tropical areas². The control of cattle ticks by chemical acaricides has become increasingly difficult due to the emergence of acaricide-resistant strains of ticks³. An apparently novel immunological method of tick control in guinea pigs and cattle is reported here. Hosts were immunised with antigens extracted from the gut and

Table 1 Results from *D. andersoni* infestations* of immunised and control guinea pigs

Guinea pigs		Mean weights ($\pm$ s.e.) of fed female ticks (mg)	Mean weights ($\pm$ s.e.) of egg mass/female (mg)	Mean numbers of larvae ($\pm$ s.e.) per female
Controls	1	297.26 ($\pm$ 63.6)	232.22 ($\pm$ 24.3)	3,575 ($\pm$ 806)
	2	401.62 ($\pm$ 56.7)	194.85 ($\pm$ 36.1)	2,709 ($\pm$ 478)
	3	396.27 ($\pm$ 72.3)	184.95 ($\pm$ 28.7)	2,973 ($\pm$ 288)
	4	318.91 ($\pm$ 96.1)	89.72 ($\pm$ 42.1)	1,366 ($\pm$ 209)
	5	31.12 ($\pm$ 2.3)	13.28 ($\pm$ 12.8)	102 ($\pm$ 73)
	6	38.66 ($\pm$ 2.8)	18.32 ($\pm$ 9.8)	0
Immunised with antigen I†	7	29.56 ($\pm$ 1.3)	0	0
	8	22.73 ($\pm$ 1.8)	0	0
	9	279.48 ($\pm$ 82.3)	47.70 ($\pm$ 36.2)	0
	10	76.71 ($\pm$ 31.8)	14.82 ($\pm$ 8.2)	0
	11	74.42 ($\pm$ 24.6)	12.08 ($\pm$ 10.9)	0
	12	146.61 ($\pm$ 76.3)	55.35 ($\pm$ 26.7)	0
Immunised with antigen II†	13	29.43 ($\pm$ 1.1)	0	0
	14	23.76 ($\pm$ 1.0)	0	0
	15	17.81 ($\pm$ 1.2)	0	0
	16	35.72 ($\pm$ 1.8)	0	0
	17	28.16 ($\pm$ 0.9)	0	0
	18	30.32 ($\pm$ 1.6)	0	0

* Each guinea pig was infested with four male and four female ticks.

† Antigen preparations: midgut, reproductive and other organs were dissected from female ticks which had fed for 5 d on guinea pigs. Dissected organs were held in ice-cold phosphate buffered saline (pH 7.2, 0.01 M in 0.15 M NaCl). Antigen I was extracted from midgut and reproductive organs, antigen II from all internal organs. Tissues were homogenised, centrifuged at 10,000g for 30 min, and the supernatants held at 4 °C. Protein concentrations were estimated by the Lowry method¹².

other internal organs of partially fed ticks. Ticks which subsequently parasitised these hosts showed significantly reduced feeding and reproductive performance. These experimental results raise the theoretical possibility of controlling some species of cattle ticks by appropriate immunisation of their hosts.

It has been known for some time that guinea pigs acquire a resistance to tick feeding⁴. For example, they allow approximately 80% of *Dermacentor andersoni* larvae to engorge during a primary 5-d infestation of one ear, but <10% of the larvae engorge during a subsequent similar infestation of the opposite ear. Such acquired tick resistance is an immunological phenomenon, abrogated by the use of immunosuppressants, and

transferable between syngeneic guinea pigs with viable lymph node cells. Tick resistance is accompanied by a cutaneous hypersensitivity to antigens derived from the salivary glands of the tick, and a degree of resistance has been induced artificially in guinea pigs using such antigens⁵⁻⁷.

Cattle may also acquire a similar resistance to the tropical cattle tick *Boophilus microplus*, following a series of infestations. However, the ability to acquire high level resistance to the tick is confined to certain strains of cattle. This ability is more often evident in *Bos indicus* than in *Bos taurus*, but the ability varies markedly from animal to animal within and between breeds⁸⁻¹¹. At present, the causes of this variability are not known, and artificial induction of high level tick resistance has not been achieved in cattle by use of tick salivary gland antigens.

Immunisation trials reported here used antigens from the ticks' midgut and reproductive organs, and were based on the crude concept that ticks feeding on appropriately immunised hosts might ingest antibodies specific for antigens within the alimentary tract and reproductive organs of the tick, producing deleterious effects on the feeding and reproductive behaviour of the tick.

Antigenic extracts were prepared from female *D. andersoni* ticks which had been allowed to feed for 5 d on outbred guinea pigs. Extracts from tissues of unfed ticks had previously been found to be ineffective. Midgut, reproductive and other organs were dissected from the ticks. Two antigenic extracts were prepared; antigen I from the midgut and reproductive organs, and antigen II from all internal organs of the tick.

Eighteen outbred guinea pigs, weighing 400–450 g, were randomly placed in three groups. Animals in group I received antigen I 1.2 mg per kg body weight in Freund's complete adjuvant (FCA) on days 0 and 14. Animals in group II were immunised with similar doses of antigen II in FCA. Control animals (group III) received similar injections, lacking tick antigens, at the same time. All injections were given intradermally into footpads. On day 28, each guinea pig was infested with four male and four female *D. andersoni*. Ticks were removed from the guinea pigs 2 weeks later. The ticks were weighed, and females were placed individually in vials held at 24 °C in a relative humidity of 85%. The egg mass produced by each female was weighed, and the numbers of larvae hatching within 35 d of removing the ticks from their hosts were recorded.

The results of this experiment are shown in Table 1. Ticks engorged on four of the six control guinea pigs and produced egg masses from which large numbers of larvae hatched. Ticks feeding on animals immunised with antigen I produced relatively few eggs from which no larvae hatched. Those placed on guinea pigs immunised with antigen II neither engorged nor laid

Table 2 Results from *D. andersoni* infestations* of immunised and control calves

		Total ticks recovered	Total weight live ticks (g)	No. of dead part-fed females	No. of females laying eggs	Total egg production ($\times 10^3$)	Total larvae produced ($\times 10^3$)
Immunised†	1	79	4.31	42	12	26.84	25.89
	2	101	2.52	40	10	10.57	0
	7	101	3.79	56	11	20.20	19.03
	8	105	6.79	38	26	52.91	47.99
Mean $\pm$ s.e.		96.5 ($\pm$ 5.2)	4.35 ($\pm$ 0.8)	44 ($\pm$ 3.5)	14.75 ($\pm$ 3.3)	27.63 ($\pm$ 7.8)	23.23 ($\pm$ 8.6)
Controls	5	115	26.55	5	57	171.66	92.98
	6	111	24.01	20	47	165.53	128.87
	10	122	31.35	15	58	224.51	183.27
	12	80	15.83	8	38	93.45	56.97
Mean $\pm$ s.e.		107 ($\pm$ 8.0)	24.44 ($\pm$ 2.9)	12 ($\pm$ 2.9)	50 ($\pm$ 4.3)	163.79 ($\pm$ 23.3)	115.52 ($\pm$ 23.3)
Significance (P value)‡			+(<0.001)	+(<0.001)	+(<0.005)	+(<0.005)	+(<0.01)

* Each calf was infested with 30 male and 100 female *D. andersoni*, confined to the ears within double cotton bags.

† Antigen I was prepared as before, from female ticks which had fed on cattle. Protein concentrations of the extracts were estimated by the Lowry method¹². Antigen preparations were held at 4 °C until used.

‡ Differences in the performances of ticks on immunised and control calves were analysed by Student's *t* test.

eggs. The success in using antigenic extracts of partially fed ticks, in contrast to previous failure with extracts of unfed ticks, may reflect the developmental and presumably antigenic changes which occur in the midgut epithelium and other tissues of the feeding female *Ixodid tick*¹. These results were considered to be sufficiently encouraging to undertake a pilot experiment with *D. andersoni* infestations of immunised and control cattle.

Antigen I was prepared, as before, from female *D. andersoni* which had fed for 5 d on the ears of Holstein steers. Eight Hereford-cross calves, which had not previously met ticks, and which weighed between 180 and 280 lb, were randomly divided into test and control groups. Each calf to be immunised received 67 mg antigen in Freund's incomplete adjuvant on days 0 and 16. A similar dose without adjuvant was administered on day 25. All injections were given by the intramuscular route. Control calves received similar injections, lacking the tick antigens, on the same days. All calves received 10-d tick infestations starting on day 28. Measurements of the ticks' feeding and reproductive performances were recorded as shown in Table 2. Serum samples were obtained from all calves on days 0, 16, 25, 28 and 38. Immunodiffusion tests were run with these samples and the immunising antigen.

From the results shown in Table 2, it is evident that there was no significant difference in the total numbers of ticks recovered from immunised and control calves. Significantly lower total weights of live ticks, lower egg production and fewer larvae were found in infestations of immunised compared with control calves. No larvae were produced by ticks which infested immunised calf number 2.

Immunodiffusion studies revealed single precipitation bands between the immunising antigen and sera collected on days 16, 25 and 28 from all immunised calves. Strong multiple bands were obtained with sera collected from calves 1, 2 and 7 on day 38, but single bands only were obtained from the serum collected from calf 8 on that day. This calf was the youngest of the group and produced the least impressive results on tick challenge.

It seems that the feeding and reproductive performances of ticks infesting immunised calves were significantly reduced. It seems reasonable to speculate that more impressive results may be obtained in cattle immunised with antigen II. Such immunisations are now being investigated. Theoretically, such an immunological method of tick control might be expected to be most effective against such species as *Boophilus microplus*, a relatively host-specific cattle tick, the larvae, nymphs and adults of which must feed in sequence from the same host. There would be potential problems in purification and large scale production of antigens but, if all cattle in an area could be successfully immunised, one might expect significant reduction in both numbers of ticks and in the tick-borne diseases of cattle.

This investigation was supported in part by a grant from the Merck Institute for Therapeutic Research, Rahway, New Jersey.

J. R. ALLEN
S. J. HUMPHREYS

Department of Veterinary Microbiology,
University of Saskatchewan,
Saskatoon, Canada

Received 4 April; accepted 27 June 1979.

1. Balashov, Y. S. *Misc. Publ. ent. Soc. Am.* **8**, 1-376 (1972).
2. Callow, L. L. *Wild Anim. Rev.* **28**, 20-25 (1978).
3. Wharton, R. H. *Wild Anim. Rev.* **20**, 8-15 (1976).
4. Trager, W. J. *Parasit.* **25**, 57-81 (1939).
5. Allen, J. R. *Int. J. Parasit.* **3**, 195-200 (1973).
6. Wikel, S. K. & Allen, J. R. *Immunology* **30**, 311-316 (1976); *Immunology* **30**, 479-484 (1976).
7. Wikel, S. K., Graham, J. E. & Allen, J. R. *Immunology* **34**, 257-263 (1978).
8. Riek, R. F. *Aust. J. agric. Res.* **13**, 532-550 (1962).
9. Roberts, J. A. *J. Parasit.* **54**, 657-661 (1968).
10. Wagland, B. M. *Aust. J. agric. Res.* **26**, 1073-1080 (1975).
11. Utech, K. B. W., Siefert, G. W. & Wharton, R. H. *Aust. J. agric. Res.* **29**, 411-422 (1978).
12. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Rundall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).

Involvement of a gene on chromosome 9 in human fibroblast interferon production

HUMAN fibroblast interferon is a single monomeric glycoprotein¹, and presumably a single structural gene codes for the human interferon polypeptide. It was originally reported that chromosomes 2 and 5 were necessary together for human interferon production². More recent evidence suggests that both chromosomes 2 and 5 carry separate human interferon genes^{3,4}. However, we have found, in two separate series of human-mouse somatic cell hybrid cell lines and their diphtheria toxin-resistant subclones^{5,6}, that human interferon production segregated from chromosome 5 and, furthermore, did not correlate with the presence of chromosomes 2 and 5 together⁷. Furthermore, using 29 hybrid clones (13 independent primary hybrids and 16 subclones) we could exclude chromosomes 1-8, 10-22 and X as candidates for the localisation of a single gene coding for human interferon on the basis of discordance of more than three clones for the segregation of human interferon production and the presence of a particular chromosome. However, there was a good correlation between the production of human interferon and the presence of chromosome 9. Here we demonstrate this directly.

The human fibroblast cell line, GM705 (Human Genetic Mutant Cell Repository, Camden, New Jersey) contains one normal inactive X chromosome and an X; 9 translocation (q12; p24) in which the long arm of X carrying the genes coding for hypoxanthine-guanine phosphoribosyl transferase (HPRT) and glucose-6-phosphate dehydrogenase (G6PD) is translocated onto chromosome 9 (refs 8, 9). When these cells are fused with mouse cells lacking HPRT, the X; 9 translocated chromosome bearing the gene for human HPRT can be selected for using the standard hypoxanthine-aminopterin-thymidine (HAT) selection procedure and selected against by 8-azaguanine. Thus, we used hybrid cells containing the X; 9 translocated chromosome to test directly the involvement of chromosome 9 in human interferon production.

Human GM705 cells, grown in Ham's F12 medium plus 20% fetal bovine serum (FBS) were fused with the HPRT mouse cell lines L-A9 (ref. 10) and PG19 (ref. 11) using polyethylene glycol (molecular weight 6,000 (ref. 12)). Hybrid cells were selected in Ham's F12 plus 10% FBS containing HAT (hypoxanthine 13.6 µg ml⁻¹, aminopterin 0.19 µg ml⁻¹, and thymidine 3.9 µg ml⁻¹) and ouabain (10⁻⁵ M)¹³, which kill mouse and human parental cells, respectively. Hybrid cell lines were isolated as previously described⁷, grown to confluency in 100-mm plastic dishes (1.5-4.0 × 10⁶ cells per dish) and induced for interferon production with Newcastle disease virus (NDV) strain F (1 × 10⁻⁴ haemagglutinating units per cell). Interferon containing fluids were collected at 20-24 h post infection, dialysed for 5 d at 4 °C against pH 2 buffer to inactivate residual NDV and returned to pH 7.4 by dialysis against culture medium. They were then assayed for human and mouse interferon in human embryo or foreskin fibroblasts and mouse L-A9 or L-929, respectively, using the method of 50% inhibition of viral ribonucleic acid synthesis (INAS₅₀)¹⁴. Neither human nor mouse interferon gave a detectable antiviral response in the respective heterologous cell system. Use of specific antisera showed that the hybrids produced human fibroblast interferon and not human leukocyte interferon.

Eight independently selected hybrid cell lines formed from the fusion of human GM705 and mouse PG19 all produced human and mouse interferons, although titres of human interferon produced were substantially less than those of mouse interferon (Table 1). GM705 fibroblasts produced 10² U per ml per 10⁶ cells when virally induced in the same conditions. Isozyme analyses of GP hybrids indicates that all eight contain enzyme markers for chromosomes X and 9 (that is, G6PD, AK1, AK3 and ACON, see Table 1 legend for abbreviations) and in

Table 1 Interferon production and human chromosomes in human-mouse cell lines

Hybrid cell line	Interferon yields*		Methods of analysis	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	X; 9				
	Human	Mouse		ENO ₁	PEPC	MDH ₅	ACP ₁	ICD ₅	—	—	HEX _β	ME ₅	—	—	AK ₁	AK ₃	ACON ₅	GOT ₅	LDH _A	PEPB	ESD	NP	MPI	APRT	GALK	PEPA	GPI	ADA	SOD _A	ACON _M	G6PD
GP-1	620	8,890	I	+	+	+	+	ND		—	+	+	+	+	ND	+	+	ND	+	+	+	+	+	+	+	+	+	+	NDND	+	
GP-2	890	178,000	I	+	+	+	+	+	+	—	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
GP-3	445	31,500	I	+	+	+	+	ND		—	+	+	+	+	ND	+	+	ND	+	+	+	+	+	+	+	+	+	+	NDND	+	
GP-4	890	22,400	I	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
GP-5	125	11,200	I	+	+	+	+	ND		—	+	+	+	+	ND	+	+	ND	+	+	+	+	+	+	+	+	+	+	NDND	+	+
GP-6	790	88,900	I	+	+	+	+	ND		—	+	+	+	+	NDND	+	+	ND	+	ND	+	ND	+	+	+	+	+	+	NDND	+	+
GP-9	500	31,600	I	ND	+	+	+	ND		—	+	+	+	+	ND	+	+	ND	+	ND	+	NDND	+	+	+	+	+	+	NDND	+	+
GP-15	620	112,000	I	+	+	+	+	+		—	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
GL-A	125	1,000	C(%)	0		36	44	0	52	60	20	60	4	0	60	44	0	28	32	48	0	60	56	52	64	0	60	64			
GL-B (Derived by diphtheria toxin selection of GL-A)	125	2,810	I	—	—	+	NDND		+	ND			see X; 9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	NDND	+	
			C(%)	0		16	40	0	0	64	36	52	0	0	56	48	0	28	32	52	0	60	52	52	60	0	64	64			
GL-C (Derived by 8- azaguanine selection of GL-B)	0	6,700	I	—	—	±	NDND		—	ND			see X; 9	+	+	+	—	+	+	+	+	+	+	+	+	+	+	+	NDND	+	
			C(%)	0		8	40	0	0	68	28	60	0	0	68	0	0	28	20	72	0	64	52	52	72	0	0	0			

Metaphase chromosomes were banded and stained using the method of Bishun *et al.*¹⁵. Twenty-five metaphase figures were analysed for each hybrid cell line and numbers of human chromosomes expressed as percent. Human isozymes used as markers for individual chromosomes were, with the exception of *N*-acetyl hexosaminidase- β , tested by routine electrophoretic methods (Harris and Hopkinson^{16,17}). *N*-acetyl hexosaminidase- β was assessed by immunodiffusion (Swallow *et al.*⁶). Abbreviations: ENO₁, enolase-1; PEPC, peptidase-C; MDH_s, soluble malate dehydrogenase; ACP₁, acid phosphatase-1; ICD_s, soluble NADP-dependent isocitrate dehydrogenase; HEX_β, *N*-acetyl hexosaminidase- β ; ME_s, soluble malic enzyme; AK₁, adenylate kinase-1; AK₃, adenylate kinase-3; ACON_s, soluble aconitase; GOT_s, soluble glutamate-oxaloacetate transaminase; LDH_A, lactate dehydrogenase-A; PEPB, peptidase-B; ESD, esterase-D; NP, purine nucleoside phosphorylase; MPI, mannose phosphate isomerase; APRT, adenine phosphoribosyl transferase; GALK, galactokinase; PEPA, peptidase-A; GPI, glucose phosphate isomerase; ADA, adenosine deaminase; SOD_A, superoxide dismutase; ACON_M, mitochondrial aconitase; G6PD, glucose-6-phosphate dehydrogenase. I, isozyme analysis; C, chromosome analysis; ND, not done.

* Titres expressed as reference research units per ml per 10⁶ cells.

† The normal inactive X chromosome was lost in GL-C, presumably due to the 'random' segregation of this chromosome.

addition have several other human chromosomes in common. The chromosome 5 marker HEX_β was absent or present only in trace amounts (Table 1). The three chromosome 2 enzyme markers tested (MDH_s, ACP₁, ICD_s) were also absent in one clone (GP-2). This particular clone (GP-2) also responds to polyribinosinic acid: polyribocytidylic acid and superinduction and has retained the capacity to produce human interferon when virally induced after 6 months in continuous serial culture in medium containing HAT.

Further hybrids were prepared from the fusion of GM705 and mouse L-A9 cells. One of these, GL-A, contained chromosome 5, was selected with diphtheria toxin at 10⁻¹ L_t ml⁻¹ (see ref. 5) in the presence of HAT, and the resulting resistant (segregant) cells grown and tested for interferon production. These derived cells, GL-B, maintained the ability to produce human interferon at the same level as the parental hybrid, GL-A (Table 1). However, when GL-B cells were exposed to 8-azaguanine (10 µg ml⁻¹) to select against the X; 9 chromosome, and resistant hybrid cells (GL-C) grown and tested for interferon production, the capacity to produce human interferon was lost (Table 1). Recently, similar results have been obtained after growth of GP-2 in 8-azaguanine. Karyotype and isozyme analyses of GL-A, B and C (Table 1) showed that chromosome 5 was eliminated by diphtheria toxin (GL-B) and X; 9T by 8-azaguanine (GL-C).

The production of high levels of human interferon by all the clones selected to contain the X; 9 chromosome, and the loss of human interferon production in the 8-azaguanine-resistant clones strongly suggest that a gene on the X; 9 chromosome is involved in human interferon production; this is presumably on 9 rather than on X, as previously we have found 11 clones discordant for the presence of X and human interferon production⁷. The only other chromosomes lost in GL-C that were present in GL-B were chromosome 12, which has been excluded in our previous study⁷, and the inactive X chromosome.

Thus, our data point to the assignment of the human fibroblast interferon gene to chromosome 9 rather than 2 or 5, as these are

absent in one or more of our hybrid clones which are inducible for human interferon. As fibroblast interferon is a glycoprotein, the products of genes other than the structural gene must be required for the secondary modification of the human interferon polypeptide. This modification may be completed by products of the rodent genome. However, we cannot exclude the possibility that genes on the other human chromosomes in addition to chromosome 9 are required for human interferon production in our hybrids, but neither 2 nor 5 seem to be involved. The basis for the discrepancy of our results from those of Tan³ and Slate and Ruddle⁴ is not clear. These authors did not observe chromosome 9 in any of their reported hybrids^{2,4}. However, Slate and Ruddle (personal communication) have now also shown that chromosome 9 does contain a structural gene for human interferon by using a human-mouse hybrid containing chromosome X; 9 as the only human chromosome. C. Chany (personal communication) also has evidence for the involvement of chromosomes 2, 5 and 9. Thus, it is possible that there is more than one structural gene for human interferon, and that they are on different chromosomes, although our evidence so far fails to show this.

We thank the MRC for a programme grant enabling us to carry out this research, and Lorraine Evans and Steve Jeremiah for technical assistance. Diphtheria toxin was from Dr D. C. Edwards (Wellcome).

A. MEAGER
H. GRAVES
D. C. BURKE

Department of Biological Sciences,
University of Warwick,
Coventry, UK

D. M. SWALLOW

MRC Human Biochemical Genetics Unit,
The Galton Laboratory,
University College London, Wolfson House,
4 Stephenson Way, London NW1, UK

Received 9 March; accepted 10 July 1979.

1. Knight, E. Jr *Proc. natn. Acad. Sci. U.S.A.* **73**, 520-523 (1976).
2. Tan, Y. H., Creagan, R. P. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **71**, 2251-2255 (1974).
3. Tan, Y. H. in *Interferons and their Actions* (ed. Stewart, W. E. II) Ch. 5 (CRC Press, Cleveland, Ohio, 1977).
4. Slatz, D. L. & Ruddle, F. H. *Cell* **16**, 171-180 (1979).
5. Creagan, R. P., Chen, S. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2237-2241 (1975).
6. Swallow, D. M., Solomon, E. & Pajunen, L. *Cytogenet. Cell Genet.* **18**, 136-148 (1977).
7. Meager, A. *et al.* *J. gen. Virol.* (in the press).
8. Pearson, P. L., Sanger, R. & Brown, J. A. *Cytogenet. Cell Genet.* **13**, 190-195 (1974).
9. Shows, T. B. & Brown, J. A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2125-2129 (1975).
10. Cox, R. P., Krauss, M. R., Balis, M. E. & Dancis, J. *Expl. Cell. Res.* **24**, 217-254 (1972).
11. Jonasson, J., Povey, S. & Harris, H. *J. Cell Sci.* **24**, 217-254 (1977).
12. Davidson, R. L. & Gerald, P. S. *Somatic Cell Genet.* **2**, 165-176 (1976).
13. Mankovitz, R., Buchwald, M. & Baker, R. M. *Cell* **3**, 221-226 (1974).
14. Atkins, G. J., Johnston, M. D., Westmacott, L. M. & Burke, D. C. *J. gen. Virol.* **25**, 381-390 (1974).
15. Bishun, N. P., Williams, D. C. & Doyle, P. in *Laboratory Manual of Cell Biology* (eds Hall, D. & Hawkins, S.) 76-78 (English Universities Press, London, 1975).
16. Harris, H. & Hopkinson, D. A. *Handbook of Enzyme Electrophoresis in Human Genetics* (North-Holland, Amsterdam, 1976).
17. Harris, H. & Hopkinson, D. A. *Handbook of Enzyme Electrophoresis in Human Genetics*, Suppl. (North-Holland, Amsterdam, 1977).

Sheep fetal haematopoietic cells produce adult haemoglobin when transplanted in the adult animal

DURING the fetal stage of development, the human erythroid cells synthesise mainly fetal haemoglobin ($\alpha_2\gamma_2$) but switch to almost exclusive formation of adult haemoglobin ($\alpha_2\beta_2$) soon after birth. Similar ontogenetic changes in the activity of the closely linked γ - and β -genes are seen in most primates and few other mammals. In goats and sheep over 95% of the haemoglobin synthesised in the fetus is of the fetal type; the switch from fetal to adult haemoglobin formation starts at about 130 gestational days and is completed within a few weeks¹. Several hypotheses have been formulated to account for these striking changes in haemoglobin gene activity during ontogeny, but the factors responsible for haemoglobin switching *in vivo* remain unknown. We report here an experiment which shows that the fetal haematopoietic cells readily switch to adult haemoglobin formation once transplanted to an adult animal. Our findings provide evidence that the switch from fetal to adult haemoglobin formation is not restricted to the perinatal developmental stage, and raise the possibility that the haematopoietic microenvironment determines the programme of haemoglobin formation in the haematopoietic progenitor cells of the fetus or the adult animal.

The experiment involved transplantation of haematopoietic cells from sheep fetuses aged 100 gestational days to adult female sheep. To have definitive markers of a successful graft of the transplanted tissue and to distinguish between the globin produced by the donor and the recipient red cells, we used the polymorphism at the β -haemoglobin locus of the sheep². There are two common alleles of this locus, β^A and β^B , directing the synthesis of the two sheep haemoglobins A ($\alpha_2\beta^A_2$) and B ($\alpha_2\beta^B_2$). The two haemoglobins can be clearly separated and quantitated by isoelectric focusing³. We therefore transplanted fetal haematopoietic cells known to be homozygous for the Hb A gene to recipient adult animals who were homozygous for the Hb B gene.

The fetal haematopoietic cells were obtained from the livers and bone marrow of twin sheep fetuses. No matching between donor and recipient cells was done. The cells (10^7 cells per kg body weight) were infused in the donor animal, which had been lethally irradiated (950 rads per total body) 2 h before the cell infusion. One of the irradiated animals died 1 d after transplantation; two other animals died at post-transplant days 11 and 15. The fourth Hb BB animal was transplanted with cells from a closely related Hb AA fetus (coefficient of kinship 0.725), received post-transplant support with leukocyte transfusions (the leukocytes were obtained from a donor sheep that

was homozygous for Hb B) and survived for 45 d; cause of death was an infection due to defective cannulation.

In this fourth animal, sequential haemoglobin analyses were carried out and findings are shown in Fig. 1. It is apparent that (1) there was a successful graft of the donor pluripotent stem cells, as Hb A appeared in the circulation of the Hb BB recipient; (2) the haematopoietic cells of the 100-d-old fetal donor switched to almost exclusive adult haemoglobin formation once they started supporting the erythropoiesis of the adult animal. There was no detectable production of Hb C by the transplanted AA cells. The small amounts of the fetal haemoglobin synthesised by the transplanted animal (Fig. 1a) were not unexpected as experiments in humans have shown that a transient production Hb F characteristically appears in post-transplantation erythropoiesis⁴. A control experiment of homo-transplantation in a BB homozygous sheep showed that in the adult sheep, as in humans, post-transplantation erythropoiesis is associated with a transient increase in fetal haemoglobin production (Fig. 1b).

In vivo studies in man have clarified certain aspects of the cellular regulation of fetal haemoglobin production. Studies in patients with clonal haematopoietic stem cell disorders suggest that a single pluripotent stem cell can provide erythroid progeny which produce fetal haemoglobin as well as progeny which produce adult haemoglobin⁵. The interpretation of this is that the regulatory events that determine Hb F or Hb A production in the adult take place in haematopoietic cells that have left the

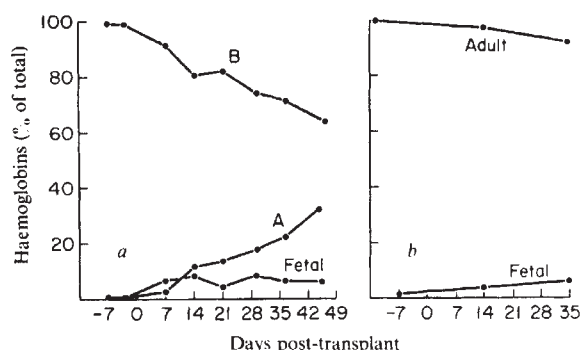


Fig. 1 Time course of haemoglobins A, B and F formation in sheep transplantation experiments. *a* Findings in a BB sheep transplanted with AA fetal sheep cells. *b* The results of a homo-transplantation experiment in a BB sheep; in *b* bone marrow cells were removed by multiple aspirations, the animal received 970 rads of whole body irradiation and was transfused with its own bone marrow cells. Note the appearance of a small amount of Hb F in the homotransplanted animal. The amount of Hb F is equivalent to that observed in experiment *a*.

pluripotent compartment of haematopoiesis. Evidence against the possibility that the pluripotent stem cells are irreversibly programmed for either fetal or adult haemoglobin expression is provided by the transplantation experiment described here. The transplanted stem cells were from a fetus which synthesised exclusively Hb F. If the Hb F phenotype of the red cell was determined at the pluripotent stem cell level, large amounts of fetal haemoglobin should have appeared in the circulation of the transplant recipient; instead, adult haemoglobin of the donor cell genotype (Hb A) appeared in the circulation as soon as there was evidence for engraftment of the transplant. Considering the size of the blood volume of the recipient and the volume of the transplanted cells, the best interpretation of the finding is that the erythroid progeny of the transplanted haematopoietic stem cells readily switched to adult haemoglobin formation.

This experiment is also informative with regard to the factors which are involved in the regulation of fetal to adult haemoglobin switching. Thus, the findings show that it is unlikely that the

temporary presence of a regulatory substance during the perinatal period of ontogeny mediates the switching from fetal to adult haemoglobin formation. It is more likely that the fetal cells were switched to adult haemoglobin formation either because they were pre-programmed to switch at a given developmental time (130–140 d from conception) or because they were re-programmed to form adult haemoglobin when they were found in the adult environment. Of the two possibilities, we consider the first one less probable. Between 100 and 135 gestational days, the cells in the sheep fetus produce predominantly Hb F (refs 1, 6); if the transplanted stem cells continued to express a pre-determined developmental programme, predominantly Hb F would have been formed during the first month of repopulation of the haematopoietic tissue of the recipient animal, but, as shown in Fig. 1a and b, the amount of Hb F in the transplanted animal was no more than that produced by the adult cells of the homotransplant control. This finding suggests that the fetal cells readily switched to an adult haemoglobin formation programme when they were moved to an adult haematopoietic environment, and raises the possibility that the change in haematopoietic environment underlays the Hb F to Hb A switching in the transplanted cells. We therefore proposed that it is the continuous presence of certain microenvironmental influences in the fetal and their absence in the adult haematopoietic tissue (or the reverse) which makes the erythroid progenitors express the fetal or the adult developmental programmes.

The phenomenon of haemoglobin switching has a twofold interest: (1) as a prototype for understanding the regulation of gene activity in development and differentiation; (2) as a therapeutic model for the treatment of sickle cell anaemia and homozygous β -thalassaemia. Patients with both these disorders could become virtually asymptomatic if it was possible to reverse the switch from Hb F to Hb A formation. If environmental factors mediate this switch, identification of these factors and their mode of the interaction with the progenitor cells may provide clues to the treatment of the β -chain haemoglobinopathies.

This work was supported by NIH grants AM 24027, CA 23021 and funds from the Veteran's Administration. G.S. is supported by NIH grants HL 20899 and GM 15253.

E. D. ZANJANI
P. B. MCGLAVER

Department of Medicine,
University of Minnesota,
Veteran's Administration Hospital,
4801 East 54th,
Minneapolis, Minnesota 55417

A. BHAKTHAVATHSALAN

Northwest Ohio Regional Perinatology Center,
Medical College of Ohio,
Toledo Hospital,
Toledo, Ohio 43606

G. STAMATOYANNOPOULOS

Division of Medical Genetics RG-20,
Department of Medicine,
University of Washington,
Seattle, Washington 98195

Received 19 March; accepted 18 June 1979.

1. Wood, W. G. *et al.* *Nature* **264**, 799–801 (1976).
2. Blunt, M. H. & Huisman, T. H. J. in *The Blood of Sheep; Composition and Function*, 155–180 (Springer, New York, 1975).
3. Suzuki, T., Benesch, R. E., Yung, S. & Benesch, R. *Analyt. Biochem.* **55**, 249–254 (1973).
4. Alter, B. P., Rapoport, J. H., Huisman, T. H. J., Schroeder, W. A. & Nathan, D. G. *Blood* **48**, 843–853 (1976).
5. Papayannopoulou, Th. & Stamatoyannopoulos, G. in *Cellular and Molecular Regulation of Hemoglobin Switching* (eds Stamatoyannopoulos, G. & Nienhuis, A. W.) 73–83 (Grune and Stratton, New York, 1979).
6. Zanjani, E. D. *et al.* in *Cellular and Molecular Regulation of Hemoglobin Switching* (eds Stamatoyannopoulos, G. & Nienhuis, A. W.) 169–177 (Grune and Stratton, New York, 1979).

Evidence of continuous evolutionary change in structures mediating adherence of lymphocytes to specialised venules

MANY cell-cell interactions are similar in a wide variety of species; for instance, penetration of ova by spermatozoa, morphogenetic cell movements during embryogenesis, and specific cellular cooperation during the immune response in vertebrates. The operation of apparently identical functional interactions in diverse species, however, does not exclude the possibility that the cell surface receptors and ligands involved may have changed considerably during evolution. In fact, studies of the molecular evolution of certain enzymes and proteins present in numerous species have demonstrated that structural evolutionary change (that is, amino acid substitutions) may proceed even in the absence of observed functional modification¹. For this reason, we suggest that the investigation of the degree of evolutionary change in structures mediating conserved cell-cell interactions requires determination of the degree to which the interaction is possible between cells from different species. The present study was designed to examine quantitatively the effect of progressive evolutionary divergence on interspecies cellular cooperation in a simple, highly conserved cell-cell interaction: the specific adherence of lymphocytes to specialised lymphocyte-binding high endothelial venules (HEV) in lymph nodes. Based on results of quantitative *in vivo* and *in vitro* assays, we report that the ability of lymphocytes from different vertebrate species to recognise specifically and bind to HEV in the mouse declines exponentially with increasing evolutionary separation of the lymphocyte donor from the mouse host. This suggests that, in spite of conservation of the functional interaction in all mammalian species examined², the cell-surface structures mediating lymphocyte adherence to HEV have been continuously and randomly altered during evolution.

As originally demonstrated by Gowans and Knight³, lymphocytes bind to and migrate through the endothelium of post-capillary HEV in lymph nodes during their normal recirculation from blood to lymph. The adherence of lymphocytes to HEV is a highly specific cell-cell interaction: other cells in the blood do not bind to HEV, and lymphocytes do not normally bind to other vessel walls. We have combined short-term *in vivo* homing studies with two previously described *in vitro* methods^{4,5} (modified in our laboratory for quantitative analysis^{6,7}) to examine the relative ability of lymphocyte populations from several vertebrate species to adhere to HEV in the mouse.

Briefly, the three assay systems used were: (1) an *in vivo* autoradiographic study, comparing the number of radiolabelled lymphocytes localising in HEV 15 min after intravenous injection; (2) a frozen section assay⁷ (modified from the method of Stamper and Woodruff⁵), comparing the binding of various lymphocyte populations to HEV in fresh frozen sections of mouse mesenteric node during incubation at 7 °C; and (3) a perfusion assay⁶ (modified from Sedgley and Ford⁴), comparing the adherence of lymphocytes to HEV during single pass systemic vascular perfusion at 37 °C. In each system, a specific adherence index (SAI) was calculated relating the observed adherence (to mouse HEV) of sample lymphocytes to that of mouse lymphocytes.

The SAI of splenic lymphocyte populations from several vertebrate species are related directly to the evolutionary distance^{8–10} separating the lymphocyte donor from the mouse host (Fig. 1a). Two conclusions can be drawn from these data. First, the *in vitro* and *in vivo* assay systems give similar results. Second, the ability of lymphocytes to adhere to host mouse HEV is inversely related to the evolutionary separation of donor and host, yet is still significantly conserved in the most divergent species tested. (It is remarkable that birds, which are separated

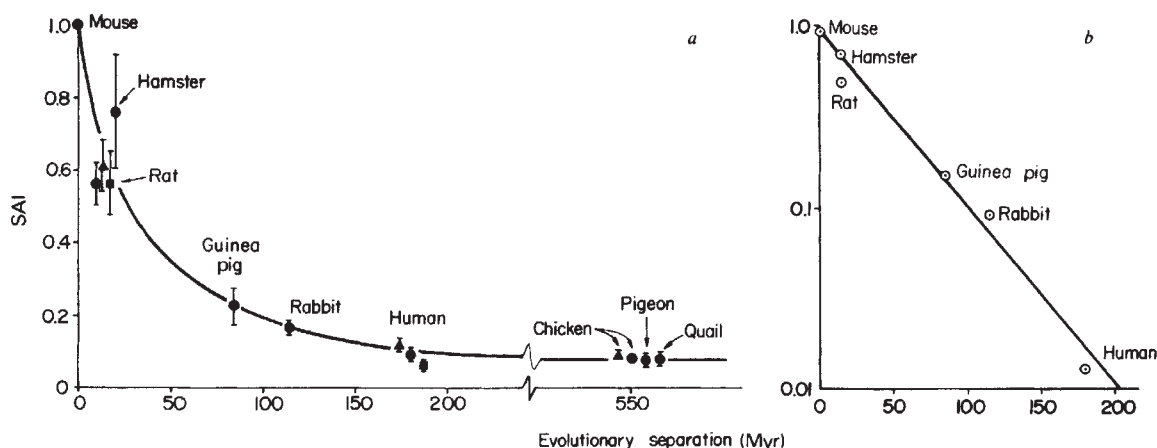


Fig. 1 *a*, Relationship between specific adherence index (SAI) and evolutionary separation of lymphocyte donor from mouse host. Adherence of sample lymphocytes to mouse lymphocyte-binding high endothelial venules (HEV) was measured *in vivo* 15 min after i.v. injection ($\blacktriangle$), during single pass systemic vascular perfusion at 37°C ($\blacksquare$), or during incubation on frozen sections of mesenteric lymph nodes at 7°C ($\bullet$). The SAI of mouse lymphocytes is unity, and that of negative control cells is 0. Error bars are standard deviations of the ratio estimates. All samples were spleen cells except rat and mouse *in vivo*, which were pooled spleen and mesenteric node cells. 5–8-week-old (BALB/c $\times$ C57BL/6 Jax)F₁ mice, Lewis rats, Syrian Gold hamsters, a 6–8-week-old Hartley guinea pig, 6–10-week-old White Leghorn or New Hampshire Red hens, 6-week and 1-year old New Zealand White rabbits, and young adult quail and King pigeons were used. Human spleens were uninvolved laparotomy specimens from young patients with Hodgkins disease. *b*, Demonstration of negative exponential relationship after subtraction of the conserved baseline adherence fraction ($K = 0.080 \pm 0.008$) typified by avian lymphocyte adherence. Data from frozen section assay. *In vivo* system ($\blacktriangle$): autoradiographs were made of sections of mesenteric lymph nodes taken from mice killed 15 min after i.v. injection of 10^8 ^3H -leucine or ^3H -uridine labelled donor lymphocytes. SAI is defined as the average number of sample lymphocytes per $\times 40$ microscope field of lymph node paracortex (either in HEV or already inside the node) divided by the corresponding value for mouse lymphocytes. Background (non-adherent labelled lymphocytes passively trapped in HEV) was less than 3% of even the most poorly adhering sample population. Frozen section assay ($\bullet$): a monolayer of lymphocytes in Hank's balanced salt solution containing 1% bovine serum albumin was incubated at 7°C for 30–45 min on freshly cut unfixed frozen sections of mouse mesenteric node. Medium was removed and the section fixed with 1% glutaraldehyde in cold phosphate-buffered saline. After fixation, non-adherent cells were gently rinsed off. To allow quantitative comparisons, an internal standard population, of fluorescent lymphocytes labelled by 15 min incubation with $60 \mu\text{g ml}^{-1}$ fluorescein isothiocyanate (FITC) at 37°C or 5–50 $\mu\text{g ml}^{-1}$ rhodamine isothiocyanate (RITC) at 25°C, was mixed with each population before incubation. A reference sample of mouse lymphocytes and a negative control population of erythrocytes or formalin-fixed lymphocytes were included. Data from at least 5 mesenteric node sections were pooled for each sample. The ratio of sample to standard cells in the incubation mixtures (R_i) and adherent to HEV (R_{HEV}) was determined. A specific adherence ratio, R_{HEV}/R_i , was calculated for each sample (SAR_s), for mouse lymphocytes (SAR_m) and for negative control cells (SAR_n). SAI was defined as: $\text{SAI} = (\text{SAR}_s - \text{SAR}_n)/(\text{SAR}_m - \text{SAR}_n)$. Standard deviations of all ratios were calculated by the delta method. Perfusion assay ($\blacksquare$): a mixture of 10^8 ^3H -leucine labelled sample spleen cells, 10^8 RITC-labelled standard cells, and 10^8 formalin-fixed FITC-labelled thymocytes (as an internal negative control) in 15 ml medium 199 containing 1% bovine serum albumin was perfused at 37°C and 2–3 ml per min through the systemic vasculature of an anaesthetised mouse (in via a trocar inserted in the left ventricle, out via an incision in the right atrium). Non-adherent cells were flushed with 10 ml medium. Solitary sample, standard, and negative control cells in HEV were counted in autoradiographs of 6- μm frozen sections of mesenteric nodes from duplicate perfusions. SAI were calculated as for the frozen section assay.

from mice by over 500 million years of evolution and lack morphologically well-developed HEV of the mammalian type², nonetheless have lymphocytes capable of recognising and adhering to mouse HEV (Fig. 2.) This decline is not limited to spleen cell populations, as a similar evolutionary decrease in adherence to mouse HEV was observed for mesenteric node lymphocytes from several mammalian species. Furthermore, in all cases examined (mouse, rat, rabbit and chicken) adherence remained a specific property of lymphocyte populations normally capable of entering lymph nodes *in vivo*: immature, non-migrating lymphocytes bound poorly in relation to their mature counterparts⁶.

As shown in Fig. 1b, the observed relationship between SAI and evolutionary distance approximates a negative exponential function with a decay rate of $2.5 \pm 0.3\%$ of the adherence index per million years. Such a clearly defined relationship requires explanation. Although the physical characteristics of cells, such as cell size, surface morphology, surface charge, membrane protein and carbohydrate content or membrane fluidity may well influence certain cellular interactions, it is difficult to invoke such nonspecific parameters in the present case. There is no correlation between lymphocyte size or surface morphology and adherence to HEV—human lymphocytes are larger and exhibit more surface microvilli than either mouse or chicken lymphocytes (when adherent to mouse HEV after vascular perfusion—E. B., van Ewijk, and I. W., unpublished). It is unlikely that either surface charge or carbohydrate content has a significant role, as adherence is unaffected by treatment of lymphocytes with neuraminidase^{11,12}. Membrane fluidity is probably not a critical parameter, as it is undoubtedly reduced at 7°C, a temperature at which lymphocytes bind well to HEV in frozen sections. A possible relationship between overall surface protein content and adherence cannot be ruled out, as light trypsinisation of the lymphocyte surface markedly reduces binding^{11,12}; it

seems more likely, however, that trypsinisation destroys a specific binding protein. Fundamentally, it is difficult to construct a role for gross physical parameters, or a mechanism for their orderly alteration during evolution, that would result in the observed negative exponential relationship between adherence and evolutionary separation.

On the other hand, it is interesting to consider this negative exponential relationship in light of the previously cited studies of the molecular evolution of proteins. Dickerson¹ has reviewed data demonstrating that several proteins (such as cytochrome *c* and haemoglobin) have incorporated amino acid substitutions at characteristic and fairly constant rates during evolution, even though they have continued to serve the same function in diverging lineages. A linear rate of amino acid substitution implies an exponential rate of decline in the fraction of mutable amino acids a protein shares with its progenitor or with diverging proteins (ignoring the minor contribution of reverse mutations back to the original amino acid). Such an exponential decline in shared amino acids could explain a roughly parallel decline in interaction compatibility of the proteins in diverging species. We propose that the observed exponential decline in lymphocyte-HEV adherence compatibility may be the functional result of a linear rate of evolutionary mutation in the components—presumably amino acids—of the cell-surface structures mediating the interaction. Although natural selection would have ensured the continuing compatibility of 'drifting' lymphocyte and HEV surface receptors within a given species, no selective pressure would have prevented progressive development of incompatibility of interaction structures in diverging lineages. Strong selection against abrupt, major changes in the interactive sites of lymphocyte and endothelial cell receptors may have protected certain essential components from mutation, thus explaining the baseline level of interspecies compatibility demonstrated by avian lymphocytes (K in Fig. 1b).

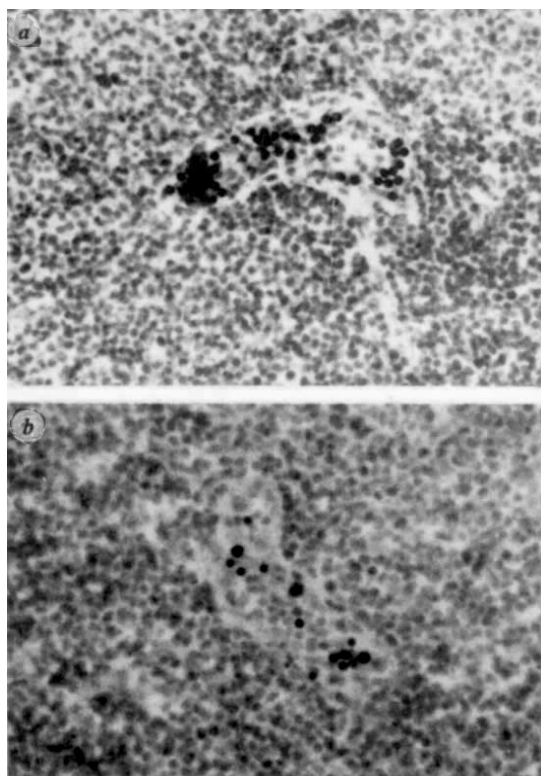


Fig. 2 Mouse (a) and chicken (b) lymphocytes (darker cells) adhering over HEV in frozen sections of mouse mesenteric lymph node ($\times 200$).

In addition to offering a plausible explanation for the observed relationship between SAI and evolutionary distance, this hypothetical model of steady evolutionary drift in interacting macromolecules predicts a similar exponential decline in interaction compatibility with increasing evolutionary separation in other systems. Examination of the ability of other simple, functionally conserved, cellular or molecular interactions to cross species barriers will test the general validity of this model. If verified, the model would imply that preservation of cellular and molecular interactions during evolution must be considered as a dynamic balance between mutation and selection, rather than static conservation.

We thank C. Wolf, J. Cockriel and S. Jacobs for secretarial and technical assistance and R. Miller for help with statistics. Supported by USPHS grant AI 09072. E.B. is supported by NIH postdoctoral training grant GM 002236-04 in Experimental Pathology, R.S. is a Special Fellow of the Leukemia Society of America, and I.W. is a Faculty Research Awardee of the American Cancer Society.

EUGENE BUTCHER
ROLAND SCOLLAY
IRVING WEISSMAN

Laboratory of Experimental Oncology,
Department of Pathology,
Stanford University Medical Center,
Stanford, California 94305

Received 1 March; accepted 10 July 1979.

1. Dickerson, R. E. *J. molec. Evol.* **1**, 26-45 (1971).
2. Miller, J. J. III *Lab. Invest.* **21**, 484-492 (1969).
3. Gowans, J. L. & Knight, E. J. *Proc. R. Soc. B* **195**, 275-282 (1964).
4. Sedgley, M. & Ford, W. L. *Cell Tissue Kinet.* **9**, 231-240 (1976).
5. Stamper, H. B., Jr & Woodruff, J. J. *J. Immun.* **119**, 772-780 (1977).
6. Butcher, E., Scollay, R. & Weissman, I. *Adv. exp. Med. Biol.* **114**, 65-72 (1979).
7. Butcher, E., Scollay, R. & Weissman, I. *J. Immun.* (in the press).
8. Young, J. Z. *The Life of Vertebrates* (Oxford University Press, New York, 1962).
9. Simpson, G. *Cold Spring Harb. Symp. quant. Biol.* **24**, 255-271 (1959).
10. Wood, A. E. in *New Encyclopedia Britannica*, Vol. 15, 976-980 (Benton, Chicago, 1973).
11. Woodruff, J. J., Katz, I. M., Lucas, L. E. & Stamper, H. B. *J. Immun.* **119**, 1603-1610 (1977).
12. Ford, W. L., Sedgley, M., Sparshott, S. M. & Smith, M. E. *Cell Tissue Kinet.* **9**, 351-61 (1976).

Distribution of fibronectin in the ectoderm of gastrulating chick embryos

A GLYCOPROTEIN known as large external transformation-sensitive (LETS) protein, cell surface protein (CSP) or fibronectin¹⁻³ can be identified on the surface of various cell types¹. This protein has a molecular weight of 210,000-250,000 and is closely related to the cold-insoluble globulin found in serum and to the factor needed for cell spreading *in vitro*^{4,5}. In culture many cells secrete fibronectin into the growth medium and on to the substratum⁶⁻⁸; fibroblasts and other cells of mesodermal origin produce particularly large amounts. However, in the early mammalian embryo fibronectin is present in basement membranes before the mesoderm has formed, when the embryo consists only of ectoderm and endoderm⁹. Adult epithelia also produce fibronectin *in vitro*¹⁰, and it forms a component of epithelial basement membranes *in vivo*¹¹⁻¹³. We have applied an immunofluorescent technique to whole mounts of early chick embryos to examine the distribution and possible morphogenetic function of fibronectin. We discuss our results here, and suggest that fibronectin is involved in morphogenetic movement during early development.

Embryos between Hamburger and Hamilton stages 3 and 6 (ref. 14) were mounted as for New culture¹⁵ and fixed in phosphate-buffered formol saline for at least 12 h. During fixation the endoderm was removed from the area pellucida and parts of the area opaca. In some cases portions of the developing mesoderm were also removed. The embryos were detached from the vitelline membrane and washed overnight with several changes of phosphate-buffered saline. Fibronectin was localised by indirect immunofluorescence^{16,17} (Fig. 1).

Fibronectin was detected in association with all parts of the ectoderm and its distribution is shown in Fig. 1. A well defined fibrous meshwork was seen around the entire perimeter of the outer area opaca margins of the blastoderm (Fig. 2) as it expanded across the vitelline membrane. The vitelline membrane itself showed no specific fluorescence. The anterior half of the boundary between the area pellucida and area opaca

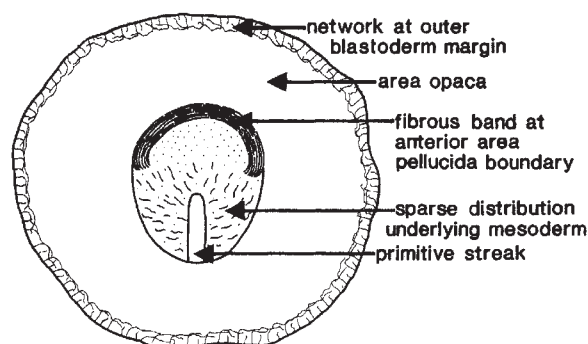


Fig. 1 Diagram to illustrate the overall distribution of fibronectin on the chick ectodermal basement membrane as detected by indirect immunofluorescence. The embryo is shown as if viewed from the ventral surface with the endoderm and mesoderm removed. Embryos were fixed in phosphate-buffered saline (PBS) containing 4% formaldehyde, and the endoderm and part of the mesoderm were removed by microdissection. They were then washed in PBS incubated at 37 °C for 30 min with rabbit anti-fibronectin antisera diluted 1:25 in PBS¹⁶, and washed and incubated for a further 30 min with fluorescein isothiocyanate-labelled goat anti-rabbit (Miles) diluted 1:10. The embryos were washed with PBS, mounted ventral side uppermost on slides in 50% glycerol, and examined in a Zeiss photomicroscope. Phase contrast optics were used to identify particular areas of the embryos and epifluorescence UV illumination to visualise the fibronectin. Photographs were taken using Ilford HP5 film (400 ASA) and the sensitivity increased to 1,600 ASA by developing with Microphen (Ilford). Exposure time was 30-120 s. A total of 23 embryos were examined. Controls included substituting the rabbit anti-fibronectin antisera with pre-immune rabbit sera or using rabbit anti-fibronectin antisera which had been passed through a fibronectin-Sepharose affinity column¹⁷. In all cases the controls showed only a low background fluorescence.

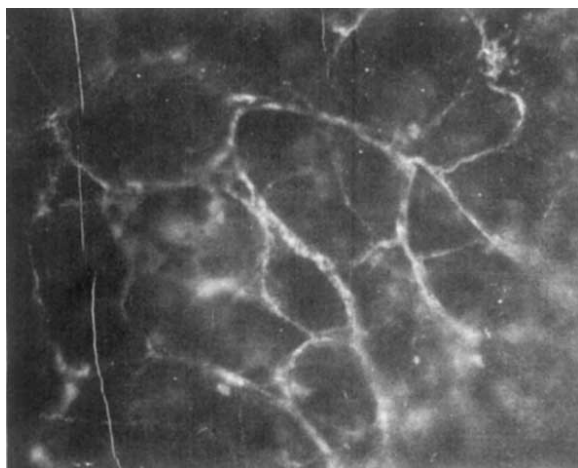


Fig. 2 Immunofluorescence demonstration of fibronectin at the outer margin of the chick blastoderm. Fluorescent fibres form a network ($\times 990$).

was outlined by an intensely fluorescent band of parallel fibres. This band was thickest in the midline anteriorly and gradually tapered as it extended posteriorly, terminating on a level with Hensen's node. Fluorescent granular deposits were present between the fibres (Fig. 3). The area pellucida on either side of the primitive streak is normally occupied by mesoderm, overlying the ectoderm. The mesoderm itself showed no specific fluorescence but if it was dissected away fibronectin could be detected on the underlying ectoderm. However, the fibronectin appeared poorly organised, the fluorescence being diffuse, with only occasional long thin fibres running mainly at right angles to the embryonic axis (Fig. 4). Between the mesoderm and the anterior boundary of the area pellucida is the proamnion, which is without mesoderm. Fluorescence there appeared as short thick disconnected rods. A very sparse and weakly fluorescent fibrous network of fibronectin was seen on the endoderm.

The fluorescence profile described above was not observed if rabbit anti-fibronectin was omitted during the staining procedure or if it was substituted with preimmune rabbit serum or rabbit antifibronectin antibodies which had previously been passed through a fibronectin-Sepharose column.

The distribution of fibronectin which we have observed in the early chick embryo can be linked to two known functions of the protein *in vitro*. First, its presence in cultures seems to facilitate cell movement¹⁸, and second, it mediates the spreading of cells on substrates¹⁹ and their attachment to collagen²⁰, although its function in the last respect may not be absolutely specific²¹. The situations in which we have observed accumulations of fibronectin in the embryo are those particularly associated with

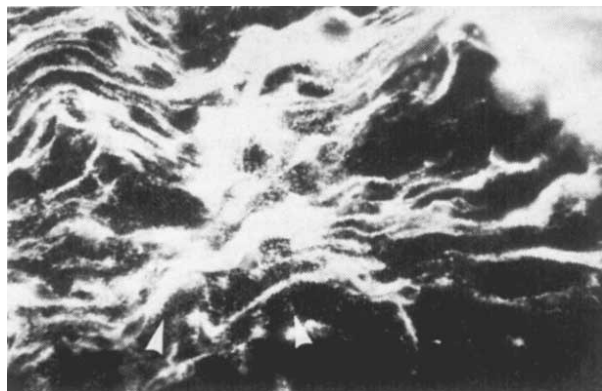


Fig. 3 Immunofluorescence photograph of part of the fibrous band at the anterior area pellucida/area opaca boundary. The fibres are intensely fluorescent, run parallel to each other and are accompanied by particulate fluorescent material (arrows) ($\times 900$).

morphogenetic movement. At the outer margins of the blastoderm the area opaca ectoderm is active in locomotion. The outer margins of the blastoderm are responsible for the expansion of the blastoderm across the vitelline membrane so that ultimately the egg yolk is enclosed in the yolk sac²². This movement is associated with a high concentration of fibronectin around the marginal ectoderm cells.

The distribution pattern of fibronectin demonstrated by immunofluorescence also coincides closely with a distribution pattern of fibres seen on the ectodermal basement membrane in the area pellucida and area pellucida/area opaca boundary zone by scanning electron microscopy (SEM)^{23,24} (Fig. 5). This pattern can be correlated with two separate morphogenetic movements, those of the mesoderm and of the primordial germ cells.

Immunofluorescence demonstrates fibronectin on the ectoderm surface underlying the mesoderm. Mesoderm cells reach their position between the ectoderm and endoderm by descending through the primitive streak and moving laterally out from the streak, across the area pellucida and eventually into the lateral and posterior area opaca, using the ectodermal basement membrane as a substrate²⁵⁻²⁷. Mesoderm cells reach the area opaca in front of the area pellucida by migrating around the anterior boundary between the area pellucida and area

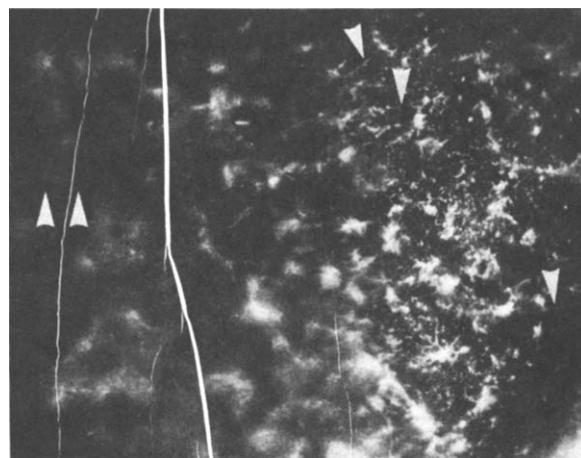


Fig. 4 An area from near the posterior end of the area pellucida. On the left side of the picture the mesoderm is only very faintly and diffusely stained. Removal of mesoderm as on the right reveals bright fluorescence, mainly finely dispersed but also as very thin fibres (arrows). The double arrow marks the embryonic axis ($\times 340$).

opaca. This boundary contains a particularly high concentration of fibronectin. The same boundary zone constitutes the germinal crescent. Primordial germ cells form there by detaching from the endoderm. The germ cells then move posteriorly around the germinal crescent to make contact with the developing vitelline vessels, reaching the gonads through the circulation^{28,29}. They have therefore moved against the main direction of mesodermal migration around the anterior area pellucida boundary. A high concentration of fibronectin in the germinal crescent may assist the germ cells in this specific migration, and in fact, SEM reveals contacts between primordial germ cells and the fibrous band in the germinal crescent²³. The developmental stage at which the fibronectin pattern is maximally defined is further evidence that its function is linked to morphogenetic movements. Mesoderm cells enter the anterior boundary zone between the area pellucida and area opaca after Hamburger and Hamilton stage 5 (refs 14, 30). Immediately before this stage the band as seen by immunofluorescence and SEM becomes most prominent²³, but after the entry of mesoderm into the boundary zone it becomes disorganised. Thus, the band has a transient existence coinciding closely with the time at which the anterior border of the area pellucida is a particularly active site of cell movement.

Fibronectin is present in the greatest quantity and in the most highly organised arrangement in the germinal crescent and

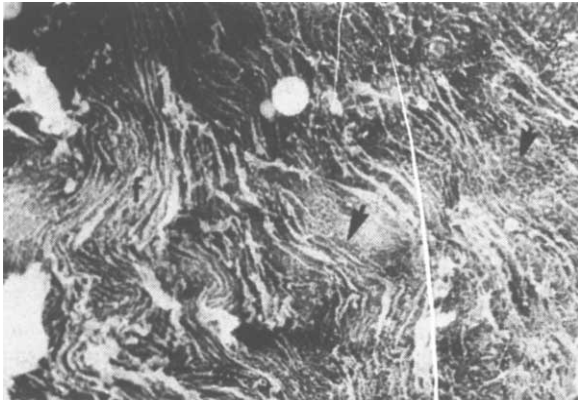


Fig. 5 Scanning electron micrograph of a portion of the fibrous band encircling the anterior boundary between area pellucida and area opaca in a stage 4 embryo. The band consists of rough-surfaced fibres in a parallel arrangement (f). They are accompanied by some loose granular material (arrows) ($\times 1,275$).

around the perimeter of the blastoderm. Both these sites are free of mesoderm cells at the time of our observations. At the blastoderm margins only ectoderm is present; in the germinal crescent the ectoderm is associated with endoderm.

We suggest that the fibronectin pattern is laid down by the ectoderm either alone or in association with endoderm and that its function is to facilitate certain morphogenetic movements in early development. First, it is involved in the overall expansion of the blastoderm across the vitelline membrane. Second, fibronectin contributes to a contact guidance system laid down by the ectoderm and utilised by primordial germ cells and mesoderm cells. The cells would tend to move along pathways of fibronectin because of the effects of the protein in favouring cell spreading and locomotion. Further work is in progress to observe these movements directly in living embryos.

We thank Miss Alison Cole for technical assistance and the Departments of Zoology and Microbiology, University of Leicester, for photomicroscopy facilities.

D. R. CRITCHLEY

Department of Biochemistry,

MARJORIE A. ENGLAND
JENNIFER WAKELY

Department of Anatomy,
University of Leicester, University Road, Leicester, UK

R. O. HYNES

Department of Biology and Centre for
Cancer Research,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

Received 19 March; accepted 5 July 1979.

- Hynes, R. O. *Biochim. biophys. Acta* **458**, 73 (1976).
- Ruoslathi, E., Vaheri, A., Kuusela, P. & Lindner, E. *Biochim. biophys. Acta* **322**, 352 (1973).
- Yamada, K. M. & Weston, J. A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3492 (1973).
- Ruoslathi, E. & Vaheri, A. *J. exp. Med.* **141**, 497 (1975).
- Grinnell, F. & Hays, D. G. *Expl Cell Res.* **115**, 221 (1978).
- Rosen, J. J. & Culp, L. A. *Expl Cell Res.* **107**, 139 (1977).
- Chen, L. B., Murray, A., Segal, J. R., Burdick, A. & Walsh, M. L. *Cell* **14**, 377 (1978).
- Vaheri, A. *et al. Ann. N.Y. Acad. Sci.* **312**, 343 (1978).
- Wariiovaara, J., Leivo, I., Virtanen, I. & Vaheri, A. *Ann. N.Y. Acad. Sci.* **312**, 132 (1978).
- Chen, L. B., Maitland, N., Galimore, P. H. & McDougall, J. K. *Expl Cell Res.* **106**, 39 (1977).
- Spiro, R. G. *Ann. N.Y. Acad. Sci.* **312**, 106 (1978).
- Linder, E., Stenman, S., Lehto, V. P. & Vaheri, A. *Ann. N.Y. Acad. Sci.* **312**, 152 (1978).
- Kefalides, N. A. *Dermatologica* **50**, 4 (1975).
- Hamburger, H. & Hamilton, H. L. *J. Morph.* **88**, 49 (1951).
- New, D. A. T. *J. Embryol. exp. Morph.* **3**, 326 (1955).
- Mautner, V. & Hynes, R. O. *J. Cell Biol.* **75**, 743 (1977).
- Hynes, R. O. & Destree, J. A. *T. Cell* **15**, 875 (1978).
- Ali, I. U. & Hynes, R. O. *Cell* **14**, 439 (1978).
- Yamada, K. M., Olden, K. J. & Pastan, I. *Ann. N.Y. Acad. Sci.* **312**, 256 (1978).
- Pearlstein, E. *Nature* **261**, 497 (1976).
- Grinnell, F. & Minter, J. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4408 (1978).

- New, D. A. T. *J. Embryol. exp. Morph.* **7**, 146 (1959).
- Wakely, J. & England, M. A. *Proc. R. Soc. B* (in the press).
- Wakely, J. & England, M. A. *J. Anat.* **128**, 435 (1979).
- Rosenquist, G. C. *Contr. Embryol. Carnegie Instn* **38**, 71 (1966).
- Trelstad, R. L., Hay, E. D. & Revel, J.-P. *Dev. Biol.* **16**, 78 (1967).
- England, M. A. & Wakely, J. *Anat. Embryol.* **150**, 291 (1977).
- Clawson, R. C. & Domm, L. V. *Am. J. Anat.* **125**, 87 (1969).
- Fujimoto, T., Ukeshima, A. & Kiyofuji, R. *Anat. Rec.* **185**, 139 (1976).
- Hamilton, H. L. *Lillie's Development of the Chick* (Holt, New York, 1952).

Insulin-receptor antibodies mimic a late insulin effect

AUTOANTIBODIES against the insulin receptor are found in patients with a rare form of extreme insulin resistance associated with acanthosis nigricans¹. These antibodies are primarily polyclonal IgGs, compete with insulin for binding to its receptor^{2,3}, immunoprecipitate solubilised insulin receptors⁴, and have been used as probes of insulin action. The acute action of the antibodies is to mimic many of insulin's biological effects, particularly those which occur rapidly after hormone binding to its cell-surface receptors⁵⁻⁹. In addition, insulin exerts biological effects which appear more slowly. For example, insulin treatment of 3T3-L1 fatty fibroblasts increases the activity of lipoprotein lipase maximally after 2-4 d but is without effect during the first 4 h (refs 10, 11). Some investigators have suggested that these slower effects may be mediated by different mechanisms from those involved in the production of the acute effects^{12,13}. We now report that autoantibodies against the insulin receptor mimic this long-term insulin effect on lipoprotein lipase.

3T3-L1 fatty fibroblasts exposed for 24 h to increasing concentrations of partially purified anti-receptor antibodies showed a dose-dependent increase in the activity of lipoprotein lipase (Fig. 1a). Significant stimulation of enzyme activity was observed with as little as 3.5 $\mu\text{g ml}^{-1}$ anti-receptor antibodies and the effect was maximal with 35 $\mu\text{g ml}^{-1}$ IgG, reaching 65% of the response induced by a maximal stimulatory insulin concentration (1 $\mu\text{g ml}^{-1}$). In contrast, normal IgG had no stimulatory effect on the activity of the enzyme. However, with very high concentrations of anti-receptor antibodies (350 $\mu\text{g ml}^{-1}$) the stimulatory activity decreased. This seems to be due to a nonspecific effect of these preparations on the cells or enzyme system, as the same high concentration of normal IgG produced a similar decrease in basal enzyme activity (Fig. 1a) and diminished the stimulatory effect of high insulin concentrations (data not shown).

As previously demonstrated, the effect of insulin on lipoprotein lipase activity in 3T3-L1 fatty fibroblasts is a slow phenomenon, requiring 12 h for half-maximal insulin stimulation^{10,11}. Similarly, the stimulatory effect of the anti-receptor antibodies is characterised by a slow time course of induction (Fig. 1b). After 12 h a maximally effective concentration of anti-receptor antibodies (35 $\mu\text{g ml}^{-1}$) was about equipotent to a maximally effective insulin concentration (1 $\mu\text{g ml}^{-1}$) in stimulating the lipoprotein lipase activity. The maximal increase in the enzyme activity in the presence of this most potent antibody concentration was observed after 24 h of exposure, and in contrast to the effect of insulin, stimulation by the antibodies then tended to decline. This is probably due to a desensitisation of the biological response at post-receptor sites observed with chronic antibody exposure, but not with insulin. We observed a similar phenomenon when studying the effect of the same antibody preparations on glucose transport and oxidation in 3T3-L1 fatty fibroblasts⁷. Thus, acute exposure of these cells to anti-receptor antibodies resulted in a stimulation of 2-deoxyglucose transport and glucose oxidation, whereas after prolonged exposure basal glucose metabolism returned to control levels. This occurred without change in the ability of the antibodies to inhibit insulin binding. Our interpretation of these data is that despite persistent occupancy of the insulin receptor the anti-receptor antibodies are unable to produce a continuing signal for generation of the biological effects.

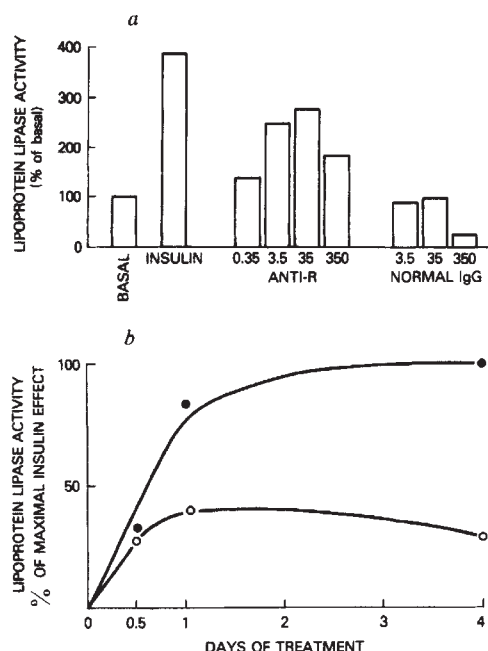


Fig. 1 *a*, Stimulation of lipoprotein lipase in 3T3-L1 fatty fibroblasts by increasing concentrations of antibodies against the receptor for insulin. 3T3-L1 cells were isolated by Dr Howard Green and provided by him and Dr Ora M. Rosen¹⁴⁻¹⁶. Fibroblasts were grown to confluence in Petri dishes with Dulbecco-Vogt's minimal essential medium containing 10% fetal calf serum. After confluence they were treated with media supplemented with 1-methyl-3-isobutylxanthine (0.5 mM) and insulin ($1 \mu\text{g ml}^{-1}$) for 2 d followed by insulin alone for 4 d to enhance conversion to fatty fibroblasts. Thereafter, cells were cultured for 4 d in Dulbecco-Vogt's minimal essential medium containing 5% fetal calf serum. The fatty fibroblasts were then exposed to insulin ($1 \mu\text{g ml}^{-1}$), partially purified antibodies against the receptor for insulin, or normal IgG. After 24 h the activity of the enzyme lipoprotein lipase was measured in the cells as described by Spooner *et al.*¹⁷. One unit of lipoprotein lipase is defined as the amount of enzyme that releases $1 \mu\text{M}$ of ^3H -oleic acid per min at 37°C from a serum-activated emulsion of ^3H -trioleoylglycerol and L- α -phosphatidylcholine. The partially purified antibodies against the receptor for insulin were prepared as follows. Plasma from a patient (patient B-2) with insulin resistance of the type B syndrome¹, containing a high titre of antibodies against the receptor for insulin (1:2,000), was treated with 33% ammonium sulphate. The precipitate was dissolved in and dialysed against 0.05 M sodium phosphate buffer, pH 7.4, containing 0.15 M NaCl, to a protein concentration of 10 mg ml^{-1} (based on absorbancy measurements assuming $A_{280 \text{ nm}}^{1\%} = 12$). This preparation is referred to as anti-R. The preparation referred to as normal IgG was obtained from plasma from a normal person using the same procedure. Results in this and all subsequent figures represent means of duplicate activity measurements each on three to six individual plates of cells. The data are expressed as per cent of basal lipoprotein lipase activity which was 0.27 ± 0.05 U per mg cell DNA (mean $\pm$ s.e.m.; $n = 4$). Lipoprotein lipase activity in the presence of insulin, 1.03 ± 0.06 U per mg DNA ($n = 4$), was statistically different from the basal activity at the $P < 0.001$ level using Student's *t*-test. Activity in the presence of 3.5 and $35 \mu\text{g ml}^{-1}$ anti-R was 0.69 ± 0.08 ($n = 4$) and 0.77 ± 0.07 U per mg DNA ($n = 4$) respectively; both values are statistically different from basal activity ($P < 0.01$). Activity in the presence of $350 \mu\text{g ml}^{-1}$ anti-R, 0.49 ± 0.05 ($n = 4$), was marginally significant ($0.05 < P < 0.10$). *b*, Time course of insulin (●) and anti-R (○) induced stimulation of lipoprotein lipase activity in 3T3-L1 fatty fibroblasts. 3T3-L1 fatty fibroblasts were exposed for 0.5, 1 and 4 d to either insulin ($1 \mu\text{g ml}^{-1}$) or antibodies against the receptor for insulin ($35 \mu\text{g ml}^{-1}$). At the end of treatment, activity of lipoprotein lipase in cells was measured and expressed as per cent of the maximal insulin stimulation above basal activity. Basal lipoprotein lipase activity was 0.33 ± 0.02 U per mg DNA ($n = 3$). The activity in cells treated with insulin and anti-R was, respectively, 0.65 ± 0.05 ($n = 3$) and 0.59 ± 0.04 U per mg DNA ($n = 3$) after 0.5 d of treatment; 1.16 ± 0.08 ($n = 3$) and 0.72 ± 0.03 U per mg DNA ($n = 3$) after 1 d; 1.33 ± 0.17 ($n = 3$) and 0.60 ± 0.02 U per mg DNA ($n = 3$) after 4 d. Stimulation of lipoprotein lipase activity by insulin and anti-R was statistically significant at all time points ($P < 0.01$).

To determine whether stimulation of lipoprotein lipase by anti-receptor antibody was mediated by mechanisms similar to those involved in the action of insulin, 3T3-L1 fatty fibroblasts were pretreated for 60 min at 37°C with anti-receptor antibodies and then exposed to increasing concentrations of insulin for 24 h at 37°C (Fig. 2). As shown above, anti-receptor antibody alone produced a 60% increase in the activity of the

enzyme. Addition of insulin at concentrations up to 10^{-6} M produced minimal further effects on the antibody-treated cells. This blockade of the insulin effect suggests that both insulin and the anti-receptor antibody use the same pathway for the production of this biological response. The small increment in the activity of the enzyme observed after the addition of high insulin concentrations is probably due to the fact that in this experiment a low antibody concentration was used, which inhibited insulin binding by only 70%.

Additional evidence that this chronic effect of anti-receptor antibodies is similar to that of insulin was obtained in studies with cycloheximide (Fig. 3). Cycloheximide, in experimental conditions in which protein synthesis was reduced by 90%, inhibited the insulin-induced increase in lipoprotein lipase activity. Similarly, the stimulatory activity of the anti-receptor antibodies on lipoprotein lipase was abolished by cycloheximide. Thus, in addition to their short-term insulin-like effects, the anti-receptor antibodies can mimic a long-term insulin effect.

The present observations have important implications for the mechanism of insulin action. Current studies have suggested two possible pathways to explain the long-term effects of insulin. First, as we have previously suggested^{9,24}, all the information required for the generation of the biological effects of insulin could be present in the receptor molecule itself. In this case, both insulin and the antibodies against the insulin receptor would generate an activated receptor which would then mediate insulin's effect either directly or by generation of the same second messenger as that for insulin. Alternatively, after binding to the cell-surface receptors, insulin and the anti-receptor antibodies are endocytosed and subsequently the ligands or their degradation product(s) bind to intracellular insulin receptors, which elicit the late insulin effect. The recent observation that insulin and/or the insulin-receptor complex enter target cells^{18,19} and the demonstration of specific binding sites for insulin on intracellular organelles²⁰⁻²³ have been taken as evidence that this last pathway may be more important.

The present results and the above-mentioned data suggest that the second hypothesis is rather unlikely. Indeed, indirect studies have indicated that the antibodies against the insulin receptor bind very poorly to intracellular receptors²⁵. Furthermore, it seems very unlikely that two totally different polypeptides, insulin and anti-receptor antibodies, would generate an identical or similar degradation product, which would

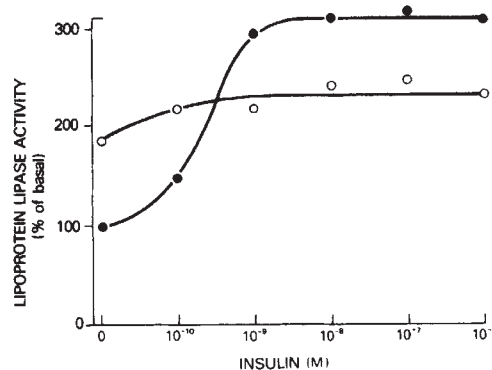


Fig. 2 Effect of antibodies against the receptor for insulin on the insulin-induced stimulation of lipoprotein lipase activity. 3T3-L1 fatty fibroblasts were pretreated for 1 h with either normal IgG ($1.75 \mu\text{g ml}^{-1}$) or anti-R ($1.75 \mu\text{g ml}^{-1}$). Increasing concentrations of insulin (0 – 10^{-6} M) were then added and the incubation continued for a further 24 h. Lipoprotein lipase activity was determined and expressed as per cent of basal. Basal lipase activity was 0.52 ± 0.13 U per mg DNA ($n = 6$). Activities in the presence of 10^{-6} M insulin alone and anti-R alone were 1.61 ± 0.10 ($n = 4$) and 0.95 ± 0.04 U per mg DNA ($n = 6$), respectively. Both values are statistically different from that of basal activity ($P < 0.001$). Lipoprotein lipase activity in cells exposed to anti-R and then 10^{-6} M insulin (1.21 ± 0.07 U per mg DNA, $n = 4$) (○), was statistically different from the activity in cells treated with 10^{-6} M insulin alone ($P < 0.05$) (●).

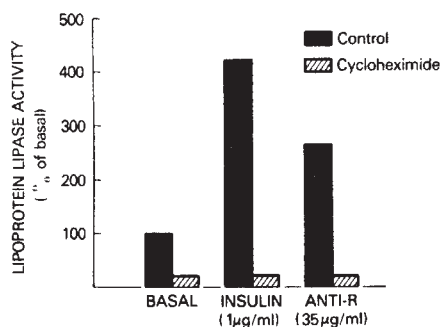


Fig. 3 Effect of cycloheximide on the lipoprotein lipase response to insulin and to antibodies against the receptor for insulin. 3T3-L1 fatty fibroblasts were incubated in basal growth medium with or without insulin ($1 \mu\text{g ml}^{-1}$) or anti-receptor antibodies ($35 \mu\text{g ml}^{-1}$) in the absence (control) or presence of cycloheximide ($4 \times 10^{-7} \text{ M}$) for 24 h. The data are expressed as per cent of the basal activity observed in control cells. Basal lipoprotein lipase activity in control cells was $0.47 \pm 0.05 \text{ U per mg DNA}$ ($n = 3$). Activity in the presence of insulin and anti-R was 1.99 ± 0.12 ($n = 3$) and $1.26 \pm 0.07 \text{ U per mg DNA}$ ($n = 3$), respectively. Both values are significantly different from that of basal activity ($P < 0.001$). In the presence of cycloheximide, lipase activity in insulin and anti-R treated cells was significantly reduced to 0.08 ± 0.01 and $0.09 \pm 0.01 \text{ U per mg DNA}$, respectively ($P < 0.001$, $n = 3$).

produce late insulin effects by binding to intracellular insulin receptors. Therefore, our present data favour the idea that both insulin and the anti-receptor antibody generate the same second messenger(s) or alternatively, that they both lead to the endocytosis of the insulin receptor, and the receptor itself contains all the necessary information for the generation of the late effects of the hormone.

Finally, it should be noted that we have dealt here only with acute and chronic 'metabolic' responses to insulin. Insulin, and several other insulin-like peptides, also produce long-term 'growth' responses, such as stimulation of thymidine incorporation into DNA. Whether these effects of insulin are mediated by a different mechanism or even the interaction of insulin with a different membrane receptor, remains to be determined.

E. VAN OBBERGHEN*†
P. M. SPOONER‡
C. R. KAHN*
S. S. CHERNICK‡
M. M. GARRISON‡
F. A. KARLSSON*
C. GRUNFELD*

* Diabetes Branch,
National Institutes of Arthritis, Metabolism and
Digestive Diseases,

‡ Section on Endocrinology,
Laboratory of Nutrition and Endocrinology,
National Institutes of Arthritis, Metabolism and
Digestive Diseases,
National Institutes of Health,
Bethesda, Maryland 20205

Received 31 January; accepted 18 June 1979.

† To whom correspondence should be addressed.

- Kahn, C. R. *et al.* *New Engl. J. Med.* **294**, 739–745 (1976).
- Flier, J. S., Kahn, C. R., Roth, J. & Bar, R. S. *Science* **190**, 63–65 (1975).
- Flier, J. S., Kahn, C. R., Jarrett, D. B. & Roth, J. *J. clin. Invest.* **58**, 1442–1449 (1976).
- Harrison, L. C., Flier, J. S., Karlsson, F. A., Kahn, C. R. & Roth, J. *J. clin. Endocr. Metab.* **48**, 59–65 (1978).
- Kahn, C. R., Baird, K. L., Flier, J. S. & Jarrett, D. B. *J. clin. Invest.* **60**, 1094–1106 (1977).
- LeMarchand-Brustel, Y., Gorden, P., Flier, J. S., Kahn, C. R. & Freychet, P. *Diabetologia* **14**, 311–318 (1978).
- Karlsson, F. A., Van Obberghen, E., Grunfeld, C. & Kahn, C. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 809–813 (1979).
- Kasuga, M. *et al.* *J. clin. Endocr. Metab.* **47**, 66–77 (1978).
- Lawrence, J. C., Larner, J., Roth, J. & Kahn, C. R. *Molec. cell. Biochem.* (in the press).

- Spooner, P. M., Chernick, S. S., Garrison, M. M. & Scow, R. O. *60th A. Mtg Endocrine Soc.* 333 (A) (1978).
- Spooner, P. M., Chernick, S. S., Garrison, M. M. & Scow, R. O. *J. biol. Chem.* **254**, 1305–1311 (1979).
- Goldfine, I. D. *Life Sci.* **23**, 2639–2648 (1978).
- Steiner, D. F. *Diabetes* **26**, 322–340 (1977).
- Green, H. & Kchinde, O. *Cell* **1**, 113–116 (1974).
- Rubin, C. S., Lai, E. & Rosen, O. M. *J. biol. Chem.* **252**, 3554–3556 (1977).
- Karlsson, F. A., Grunfeld, C., Kahn, C. R. & Roth, J. *Endocrinology* **104**, 1383–1392 (1979).
- Spooner, P. M., Chernick, S. S., Garrison, M. M. & Scow, R. O. *J. biol. Chem.* (in the press).
- Gorden, P., Carpentier, J. L., Freychet, P., Le Cam, A. & Orci, L. *Science* **200**, 782–785 (1978).
- Kahn, C. R. & Baird, K. J. *J. biol. Chem.* **253**, 4900–4906 (1978).
- Horvat, A. L. E. & Katsoyannis, P. G. *Biochim. biophys. Acta* **382**, 609–620 (1975).
- Goldfine, I. D. & Smith, G. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1427–1431 (1976).
- Kahn, C. R. *J. Cell Biol.* **71**, 261–286 (1976).
- Bergeron, J. J. M., Evans, W. H. & Geschwind, I. I. *J. Cell Biol.* **59**, 771–776 (1973).
- Kahn, C. R., Baird, I. L., Jarrett, D. B. & Flier, J. S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4209–4213 (1978).
- Goldfine, I. D., Vigneri, R., Cohen, D., Plam, N. B. & Kahn, C. R. *Nature* **269**, 698–700 (1977).

β -Adrenoreceptors determine affinity but not intrinsic activity of adenylate cyclase stimulants

SIGNIFICANT progress has recently been made toward defining and understanding the molecular basis of hormone sensitivity in the β -adrenoreceptor-adenylate cyclase system. Through the use of biochemical^{1,2} and genetic^{3,4} techniques it has become apparent that this system is comprised of at least three separate components: the β -adrenoreceptor, which is responsible for the binding of β -adrenergic agents; the catalytic unit of the enzyme adenylate cyclase, responsible for the conversion of ATP to cyclic AMP; and a nucleotide regulatory component or components which seems to possess a GTPase activity⁵ and to be required for coupling of the β -adrenoreceptor to the adenylate cyclase enzyme⁶. As has been demonstrated by Orly and Schramm⁷ and confirmed by Schwarzmeier and Gilman⁸ in cell fusion experiments, these components interact to produce the hormone-responsive adenylate cyclase system and this interaction occurs even when the separate functions are donated by heterologous cell types. Frog and turkey erythrocytes each possess a catecholamine-responsive adenylate cyclase. Although similar in some respects these two well-studied systems differ markedly with respect to their drug affinities and intrinsic activities. Intrinsic activity is defined as the maximal ability of a drug to stimulate a biological or biochemical response. To elucidate the role of the β -adrenoreceptor in the expression of these properties, we report here the use of cell fusion techniques to construct a hybrid which possesses the β -adrenoreceptor of the turkey erythrocyte and the adenylate cyclase of the frog erythrocyte. The characteristics of this hybrid indicate that although the receptor determines relative drug affinities (for example, β_1 compared with β_2 subtypes of receptor), it is a component distal to the receptor which determines relative drug intrinsic activities.

Hybrid cells were constructed by fusing frog erythrocytes, the β -adrenoreceptor of which had been inactivated with dicyclohexyl carbodiimide, with turkey erythrocytes, the adenylate cyclase of which had been inactivated by treatment with *N*-ethyl maleimide (Fig. 1 legend). Under the conditions used to inactivate the β -adrenoreceptor of the frog erythrocyte the ability of catecholamines to stimulate adenylate cyclase is totally abolished, whereas only 25–30% of basal, fluoride-, and Gpp(NH)p-stimulated adenylate cyclase activity is lost. The fact that the adenylate cyclase in these treated cells remains responsive to Gpp(NH)p indicates the functional presence of the nucleotide regulatory component as well as the catalytic unit of adenylate cyclase.

In the turkey erythrocyte treatment with *N*-ethyl maleimide results in >99% inactivation of control adenylate cyclase activity. However, when assayed for catecholamine-stimulated GTPase activity according to the method of Cassel and

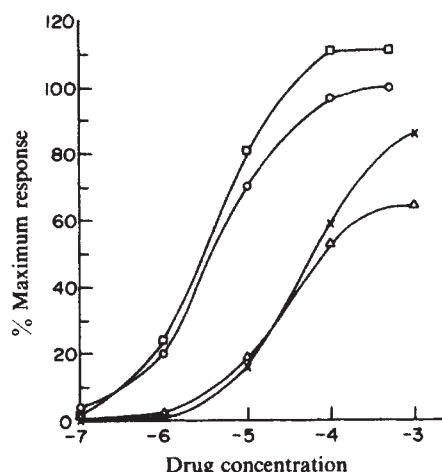


Fig. 1 Dose-response curves of agonists for stimulation of adenylate cyclase in hybrid cell membranes. $\square$, Hydroxybenzylisoproterenol; $\circ$, isoproterenol; $\times$, adrenaline; Δ , noradrenaline. The β -adrenoreceptor of the frog erythrocyte was inactivated by treatment of whole cells with 50 M dicyclohexyl carbodiimide (DCCD) for 30 min at 30°C as described previously⁷. The adenylate cyclase of the turkey erythrocyte was inactivated by treatment of the whole cells with 5 mM *N*-ethyl maleimide (NEM) for 30 min at 0°C, followed by three washes with buffered saline. Sendai virus was grown according to Watkins¹⁰. For cell fusion 1×10^9 NEM-treated turkey cells and 2×10^8 DCCD-treated frog cells were mixed in 10 ml buffer containing: 135 mM KCl, 5 mM NaCl, 0.8 mM $MgCl_2$, 2 mM $MnCl_2$ and 20 mM Tris-HCl at pH 7.4. Approximately 5–10,000 haemagglutinating units of inactivated Sendai virus was added to the cell suspension which was incubated on ice for 10 min and then transferred to a 37°C water bath for 30–60 min. Normally after microscopical examination of an aliquot of these cells, 50–75% of the nuclei were in polykaryons with greater than 95% of the giant cells containing at least one nucleus (determined by relative nucleus size) of each cell type. Membranes were prepared using a freeze-thaw lysis technique¹¹ and adenylate cyclase activity was assayed as described previously^{12,13} with the exception that the final assay mixture contained 0.3 mM ATP and 0.1 mM GTP. For each fusion experiment a 'mixed control' was run in which the appropriate number of DCCD-treated frog and NEM-treated turkey cells were mixed, incubated in fusion buffer without virus, and assayed for adenylate cyclase activity. In no case was isoproterenol-stimulated enzyme activity observed in this preparation. In addition, homologous control fusions of DCCD-treated frog cells and NEM-treated turkey cells did not result in the appearance of hormone-sensitive adenylate cyclase, nor did fusion of wild type frog erythrocytes or wild type turkey erythrocytes with themselves alter the patterns of agonist affinities or intrinsic activities. Maximum response refers to adenylate cyclase activity in the presence of 10^{-4} M isoproterenol and represents an average of 118.3 pmol cyclicAMP per mg protein per 10 min over a basal activity of 45.1 pmol per mg protein per 10 min. The EC_{50} values for the agonists for the activation of adenylate cyclase are: hydroxybenzylisoproterenol 4×10^{-6} M; isoproterenol 4×10^{-6} M; adrenaline 4×10^{-5} M; noradrenaline 3×10^{-5} M. EC_{50} is defined as the concentration of agonist required to elicit 50% of the maximal stimulation of adenylate cyclase produced by that particular drug. The curves shown are the average of four separate experiments.

Sellinger¹⁴, membranes derived from these treated cells exhibited only a 15–20% loss of catecholamine-stimulated GTPase activity. Thus, these cells seem to possess both the β -adrenoreceptor and the nucleotide regulatory protein.

In the turkey erythrocyte catecholamine agonists demonstrate a β_1 order of potency (isoproterenol > noradrenaline = adrenaline) for the activation of adenylate cyclase¹⁵. In the frog erythrocyte, however, a β_2 pattern (isoproterenol > adrenaline > noradrenaline) of catecholamine agonist potencies is observed¹⁶. This definition of β -adrenoreceptor subtypes has been set forth by Lands¹⁷. Shown in Fig. 1 are the dose-response curves of four β -adrenoreceptor agonists for stimulating

adenylate cyclase activity in membranes derived from hybrid cells. The preparation demonstrates a typical β_1 adrenoreceptor specificity for these agonists. Isoproterenol is more potent than noradrenaline and adrenaline which are approximately equipotent. In addition, hydroxybenzylisoproterenol, a β_2 -specific agonist, which is tenfold more potent than isoproterenol in the frog β_2 system¹⁸ exhibits the β_1 characteristic of equipotency with isoproterenol in the hybrid cells (L. J. Pike and R. J. Lefkowitz, unpublished results).

As the parental turkey erythrocyte adenylate cyclase and the frog erythrocyte β -adrenoreceptor have been completely inactivated, the observation that β -agonists can stimulate adenylate cyclase activity in this system suggests that a hybrid receptor-adenylate cyclase has been constructed. The β_1 specificity of the agonist response indicates that the relative affinities of β -adrenergic agents are determined by the receptor protein itself.

To determine whether the receptor was also responsible for determining the intrinsic activities exhibited by agonist drugs, the maximal ability of a series of adrenergic drugs to stimulate the adenylate cyclase in hybrid membrane preparations was examined. The results of these studies are shown in Table 1.

It can be observed that in all cases the intrinsic activities in the hybrid closely resemble those of the frog system. Particularly striking differences between agonist intrinsic activities in the turkey and frog erythrocyte are seen for noradrenaline, cobefrin and isoetharine. In each of these instances the hybrid system's response is similar to that of the frog. For all the drugs tested, the hybrid intrinsic activities are not significantly different from the frog erythrocyte intrinsic activities (on the basis of Student's *t* tests), whereas the hybrid intrinsic activities are significantly different ($P < 0.01$) from those of the turkey erythrocyte. As the intrinsic activities parallel those of the adenylate cyclase donor rather than the receptor donor, it seems that some component distal to the receptor is responsible for determining intrinsic activity.

In our current conceptualisation of the system this distal component might be either the nucleotide regulatory component, catalytic unit of adenylate cyclase, or both. It is evident that only the frog erythrocyte catalytic adenylate cyclase unit is active in the hybrids. However, the hybrids seem to contain both turkey and frog nucleotide components. Thus, it is unclear what role the nucleotide regulatory component(s) might play in the determination of intrinsic activity, as one or both might be functioning in the hybrid.

Table 1 Intrinsic activities of β -adrenergic agents in membrane preparations of frog, turkey and hybrid erythrocytes

Agonist	Intrinsic activity		
	Hybrid	Frog	Turkey
(–)Isoproterenol	1.00 (9)	1.00 (4)	1.00 (4)
(–)Adrenaline	0.87 ± 0.03 (9)	0.95 ± 0.02 (3)	1.00 ± 0.01 (3)
(–)Noradrenaline	0.62 ± 0.03 (9)	0.54 ± 0.02 (4)	1.00 ± 0.02 (4)
(±)Cobefrin	0.50 ± 0.01 (4)	0.47 ± 0.03 (4)	0.92 ± 0.03 (4)
(±)Isoetharine	0.35 ± 0.04 (5)	0.42 ± 0.01 (3)	0.71 ± 0.02 (3)
(–)Soterolol	0.13 ± 0.02 (5)	0.11 ± 0.01 (3)	0.44 ± 0.04 (3)
Dobutamine	0.07 ± 0.01 (5)	0.07 ± 0.01 (3)	0.19 ± 0.01 (3)

Control frog and turkey intrinsic activities were assessed using fused frog and fused turkey erythrocytes prepared in the same manner as the hybrid erythrocytes. To determine intrinsic activities several concentrations of each drug over a 20-fold range bracketing the saturation point were assayed for stimulation of adenylate cyclase activity. The highest activity observed (minus basal) for each was divided by the maximal activity (minus basal) observed in the presence of saturating concentrations of isoproterenol and the values are reported as this fraction. Activities were (pmol cyclicAMP per mg protein per 10 min): basal 183; isoproterenol 3108 in the frog; basal 7.6; isoproterenol 99.1 in the turkey; and basal 45.1; isoproterenol 118.3 in the hybrids. The values given are the mean $\pm$ s.e.m. of the number of experiments shown in parentheses.

The construction of a hormone-sensitive adenylate cyclase system containing components from the frog and turkey erythrocyte systems has allowed us to assess the possible role of the β -adrenoreceptor in the determination of the affinities and intrinsic activities of drugs. As has been previously observed by others^{7,8,19} we find that the receptor molecule seems to confer β -adrenoreceptor specificity to an adenylate cyclase system. In addition, however, the receptor also determines the relative affinities of β -adrenergic agents. Thus, the specificity observed in the β_1/β_2 subtypes most likely arises from slight differences in receptor protein structure. Interestingly, however, these data indicate that it is not the receptor but rather some component(s) distal to it that is responsible for the determination of intrinsic activity. Characterisation of the reciprocal hybrid (*N*-ethyl maleimide frog cell crossed with dicyclohexyl carbodiimide treated turkey cell) would further elucidate this point. However, to date all attempts to construct a functional hybrid of this composition have been unsuccessful.

The reconstitution of a catecholamine-stimulated adenylate cyclase system using cell fusion of these two different but complementary cell types is a potentially powerful tool that will be useful in defining the molecular components responsible for a variety of system characteristics such as desensitisation and the differing patterns of nucleotide regulation of adenylate cyclase activation.

This work was supported by grants from the HEW (nos HL16037 and HL20339) and by a grant-in-aid from the American Heart Association.

LINDA JOY PIKE
LEE E. LIMBIRD
ROBERT J. LEFKOWITZ

Howard Hughes Medical Institute Laboratory,
Departments of Medicine and Biochemistry,
Duke University Medical Center,
Durham, North Carolina 27710

Received 15 February; accepted 13 June 1979.

1. Limbird, L. E. & Lefkowitz, R. J. *J. biol. Chem.* **252**, 799–802 (1977).
2. Haga, T., Haga, K. & Gilman, A. G. *J. biol. Chem.* **252**, 5776–5782 (1977).
3. Insel, P. A. *et al. Molec. Pharmacol.* **12**, 1062–1069 (1976).
4. Ross, E. M. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3715–3719 (1977).
5. Cassel, D. & Selinger, Z. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3307–3311 (1977).
6. Haga, T., Ross, E. M., Anderson, H. J. & Gilman, A. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2016–2020 (1977).
7. Orly, J. & Schramm, M. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4410–4414 (1976).
8. Schwarzmeier, J. D. & Gilman, A. G. *J. Cyclic Nucleotide Res.* **3**, 227–238 (1977).
9. Pike, L. J. & Lefkowitz, R. J. *J. Cyclic Nucleotide Res.* **4**, 27–34 (1978).
10. Watkins, J. F. *Meth. Virol.* **5**, 1–32 (1971).
11. Mickey, J. V., Tate, R., Mullikin, D. & Lefkowitz, R. J. *Molec. Pharmacol.* **12**, 409–419 (1976).
12. Lefkowitz, R. J. *J. biol. Chem.* **249**, 6119–6124 (1974).
13. Salomon, Y., Londos, C. & Rodbell, M. *Analyt. Biochem.* **58**, 541–548 (1974).
14. Cassel, D. & Selinger, Z. *Biochem. biophys. Res. Commun.* **77**, 868–873 (1977).
15. Brown, E. M., Fedak, S. A., Woodard, C. J., Aurbach, G. D. & Rodbard, D. *J. biol. Chem.* **251**, 1239–1246 (1976).
16. Mukherjee, C., Caron, M. G., Mullikin, D. & Lefkowitz, R. J. *Molec. Pharmacol.* **12**, 16–31 (1976).
17. Lands, A. M., Arnold, A., McAuliffe, J. P., Luduena, F. P. & Brown, T. C. *Nature* **214**, 597–598 (1967).
18. Lefkowitz, R. J. & Williams, L. T. *Proc. natn. Acad. Sci. U.S.A.* **74**, 515–519 (1977).
19. Schramm, M., Orly, J., Eimerl, S. & Korner, M. *Nature* **268**, 310–313 (1977).

Acetylcholinesterase like that of skeletal muscle in smooth muscle reinnervated by a motor nerve

SKELETAL muscles contain several molecular forms of acetylcholinesterase (AChE) characterised by different sedimentation coefficients^{1–4}. In rat muscle^{1,2} (but not human³ and chicken⁵ muscle) the heavy form of AChE is exclusively present at the motor endplate regions. In contrast, the heavy form of AChE has not been detected in cat⁶ or rat² smooth muscles. The endplate AChE of rat skeletal muscle is selectively detached by treatment with collagenase^{1,7}, an enzyme which has been shown specifically to digest the collagen-like tail of the heavy form of

AChE from several sources^{8,9}. The presence of endplate AChE is neurally regulated¹⁰, and there is evidence that this effect is partially mediated by muscle activity¹¹ and by a trophic factor conveyed by axoplasmic transport^{12,13}. Previous results from our laboratory have shown that reinnervation of the cat nictitating membrane smooth muscle by a motor nerve results in an increase of AChE activity¹⁴. We report here that collagenase treatment detaches active AChE from cat smooth muscle reinnervated by a cholinergic motor nerve, and that a new AChE form with a high sedimentation coefficient is induced in these muscles. Our results suggest that the presence of the motor nerve determines, in reinnervated smooth muscle, the appearance of an AChE similar to that found in skeletal muscle.

The superior cervical sympathetic ganglion was removed from adult cats, the hypoglossal nerve was sectioned and the central stump sutured to the postganglionic fibres of the superior cervical ganglion. Details of the experimental procedure are given elsewhere¹⁴. Seven months post-operation, both the hypoglossal and the preganglionic nerves were stimulated to confirm that smooth muscles had been reinnervated exclusively by the hypoglossal. In all cases, the nictitating membrane and the pupil responded only to hypoglossal stimulation. The inferior smooth muscles of both nictitating membranes were then removed for AChE assay¹⁵. In one series, control and reinnervated muscles were incubated *in vitro* at 37°C in a medium¹⁶ containing 0.2 mg ml⁻¹ collagenase (preparation type III, 400 U mg⁻¹, Sigma). We used the collagenase digestion technique which is known to detach selectively AChE from the motor endplates, leaving the non-endplate AChE intact⁷. In another series of experiments, homogenates of reinnervated and normal smooth muscle were sedimented in a linear (5–20%) sucrose gradient following the method of Martin and Ames¹⁷. Sucrose solutions (Schwarz-Mann, special enzyme grade) containing NaCl-Triton X-100 were used¹. The specific activity of AChE in control muscles was 0.57 nmol ³H-acetate formed per min per mg protein, while in reinnervated muscles it increased to 1.23. Smooth muscles reinnervated for a longer period (10 months) showed a higher AChE level, reaching a specific activity of up to 3.1 (ref. 18).

Reinnervated smooth muscles incubated with collagenase released an amount of AChE proportional to the time of incubation whereas control muscles showed no solubilised AChE activity (Fig. 1). It is known that collagenase disrupts the muscle fibre basement membrane^{19,20}, and releases motor endplate AChE which is partly localised within the basal lamina at the neuromuscular junction²¹. The enzymatic detachment of active AChE from reinnervated smooth muscle suggests that this new enzyme form is probably anchored to a basal lamina-like material—as has been proposed for AChE forms which present sensitivity to collagenase⁸—and would correspond to a high sedimentation coefficient form, similar to the AChE form specific for rat motor endplate¹. The AChE release experiments, however, do not yield definite structural information since commercial collagenases are contaminated with several proteases²². The molecular forms of AChE present in reinnervated smooth muscle were therefore identified by velocity sedimentation analysis. Figure 2 shows that a new heavy form of AChE appears in smooth muscle reinnervated by a cholinergic motor nerve. This heavy form of AChE which probably corresponds to the asymmetric form A₁₂ of Bon *et al.*⁵, accounts for 9% of the total activity and represents ~25% of the total AChE increase in reinnervated muscles. As the heavy form of AChE is not present in normal smooth muscle⁶, it is likely that the species detached by collagenase corresponds to this high sedimentation coefficient form²³. The present results show that a motor nerve can induce in smooth muscle a form of AChE similar to that found in skeletal muscle. As innervation by a cholinergic motor nerve is a common factor for reinnervated smooth muscle and skeletal muscle, our results suggest that the motor nerve is able to regulate AChE in any musculature it innervates.

Several different mechanisms could be involved in neural AChE regulation: (1) muscle activity; (2) acetylcholine itself; (3)

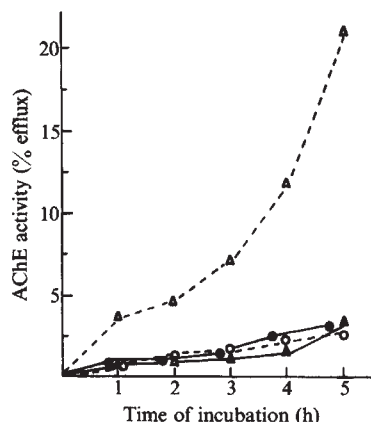


Fig. 1 Release of AChE from inferior smooth muscles of the nictitating membrane as a function of time of incubation. Experimental conditions: normal smooth muscle of the nictitating membrane with collagenase (○) and without collagenase (●), reinnervated smooth muscle of the nictitating membrane with collagenase (△) and without collagenase (▲). Smooth muscle samples (15–20 mg) were incubated at 37 °C in 300 µl of a medium containing 138 mM NaCl, 4 mM KCl, 1 mM MgSO₄, 50 mM sodium phosphate (pH 7.5), 0.1 mg ml⁻¹ chloramphenicol, and 0 or 5 µM BW 284C51 dibromide. Collagenase (0.2 mg ml⁻¹) was added to the incubation medium of some of the muscle samples. At various times, duplicate aliquots of 10 µl incubation medium were removed for AChE assay¹⁵. Subsequently, the muscle samples were homogenised 1/15 (w/v) in ice-cold buffer pH 7.3 (1 M NaCl, 0.1% Triton X-100, 0.05 M Tris-HCl, 0.2 mM EDTA) to determine AChE activity. Untreated muscle samples were also homogenised and assayed for AChE. The absolute amount of AChE in incubated samples is not the same for the different experimental situations studied, therefore to evaluate comparatively the AChE release in each case¹⁶ the initial amount of enzyme present in the samples must be taken into account. % AChE efflux:

$$\frac{(\text{AChE medium} \times 100\%)}{(\text{AChE medium} + \text{AChE in homogenate})}$$

release of AChE from the nerve terminal; and (4) neurotrophic control. These four mechanisms are discussed below:

(1) Muscle activity has been shown to play a role in modulating AChE in chick muscle culture²⁴, and stimulation of muscular activity can induce AChE spots in rat muscle¹¹. The 16S AChE form reappears at the original endplates in denervated muscles, after reinnervation at a different site²⁵. A possible explanation for this is that muscle activity stimulates synthesis or stabilisation of 16S AChE at sites on the muscle previously determined by nerve contact^{25,26}. Lømo and Slater¹¹ suggest that muscle activity might turn on synthesis of junctional AChE in addition to turning off synthesis of extrajunctional acetylcholine receptors in rat skeletal muscle. We know that reinnervation of the nictitating membrane smooth muscle by a motor cholinergic nerve determines both a decrease in the hypersensitivity to acetylcholine²⁷, and the induction of a heavy form of AChE; hence, the effect suggested for muscle activity in skeletal muscle¹¹ might also apply in smooth muscle reinnervated by a motor nerve. Recently, however, it has been shown that, in the rat, the endplate 16S AChE form is influenced by the length of the nerve stump²⁸. As muscles become inactive at the time of nerve section²⁹ regardless of the nerve stump length, it follows that the endplate 16S AChE form also depends on an activity-independent mechanism²⁸.

(2) Acetylcholine *per se* does not seem to be an important factor in AChE regulation: during reinnervation of skeletal muscle with preganglionic parasympathetic cholinergic fibres, AChE decreased and almost disappeared when reinnervation was completed, despite normal acetylcholine release³⁰.

(3) Release of neural AChE has been demonstrated in rat phrenic nerve-hemidiaphragm³¹. Although the motor nerve contains a heavy form of AChE^{32,33}, only the 10S form of AChE

has been shown to be released from nerve terminals. Evidence that a heavy form of AChE is conveyed by axonal flow was obtained by studying the accumulation of material at the proximal stump of a damaged nerve³², so it is possible that the heavy form of AChE is produced locally either in Schwann or other extra-axonal cells, at the site of the injured stump. In this context, it has been recently demonstrated that glial cell acid phosphatase is transported to damaged axons³⁴. In any case, as the axonal flow is an important mediator in the neurotrophic regulation of junctional AChE¹², the motor nerve innervating the smooth muscle of the nictitating membrane could provide a source for the heavy form of AChE.

(4) AChE trophic substances^{35,36} are conveyed by axoplasmic transport^{12,13} and in motor nerves a trophic factor has been detected which is capable of maintaining the level of AChE in cultures of aneural embryonic muscle or denervated adult chicken muscle³⁶. Moreover, a trophic factor, other than AChE, isolated from rat phrenic nerve and transported by axonal flow, has been shown to maintain the levels of endplate AChE in rat diaphragm tissue culture¹³. This finding strongly suggests that a substance different from AChE, conveyed by axonal flow to the nerve terminal, would be a major factor in the neurotrophic regulation of endplate AChE. The effect we observed in reinnervated smooth muscle may be partly mediated by a similar factor. Thus motor nerves regulate AChE, but parasympathetic

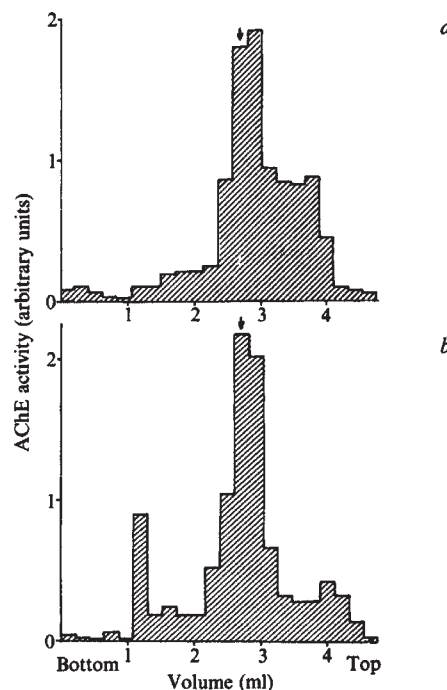


Fig. 2 Velocity sedimentation of smooth muscle extracts in a sucrose gradient. AChE activity profile of: *a*, control smooth muscle; *b*, reinnervated smooth muscle. Smooth muscles (~10 mg) were homogenised in ice-cold NaCl-Triton buffer, pH 7.3 (1M NaCl, 0.5% Triton X-100, 0.05M Tris-HCl, 0.2 mM EDTA). After various times of enzyme extraction, samples of 0.2 ml (2.4 mg protein) were layered on linear (5–20%) sucrose gradients, in the presence of the same buffer as for homogenisation and subjected to velocity sedimentation. Ultracentrifugation was performed in a Beckman L5-65 ultracentrifuge, with a SW 65 rotor at 39,000 r.p.m. for 12 h at 20 °C. A volume of 0.05 ml of each fraction was assayed for AChE activity. Activity was measured by a radiochemical procedure in which release of ³H-acetate from ³H-ACh iodide (NEN, 50 mCi mmol⁻¹) was quantified by ion-exchange chromatography and scintillation spectrometry^{15,16}. AChE activity is expressed in arbitrary units. In our standard assay conditions the hydrolysis of acetylcholine was inhibited by 80% with 5 µM BW 284C51 dibromide, a specific inhibitor of AChE activity. Catalase, used as a marker enzyme (11.3S), was determined colorimetrically³⁸ and proteins by the method of Lowry *et al.*³⁹. The arrow represents the position of catalase.

nerve³⁰ do not possess this capacity. In addition, motor nerves can induce high sedimentation coefficient AChE, similar to the specific form of rat motor endplates, as has been reported for denervated muscle endplates², non-synaptic segments of skeletal muscles², original endplates in denervated muscles after reinnervation at a different site²⁵, primary cultures of rat muscle cells³⁷, and smooth muscle (this paper).

We thank Drs H. L. Fernandez, J. Koenig, T. Lømo and J. Massoulié for making their results available before publication. This research was supported by grants from UNDP/UNESCO, DIUC and Gildemeister Foundation.

NIBALDO C. INESTROSA*
BERNARDITA MENDEZ†
JOAQUIN V. LUCO

Laboratory of Neurophysiology,
Department of Cell Biology,
Catholic University,
Santiago, Chile

Received 26 February; accepted 20 June 1979.

* Present address: Laboratory of Biochemical Cytology, Department of Cell Biology, Catholic University, Santiago, Chile.

† Present address: Endocrine Research Division, University of California, San Francisco, California 94143.

- Hall, Z. W. *J. Neurobiol.* **4**, 343–361 (1973).
- Vigny, M., Koenig, J. & Rieger, F. *J. Neurochem.* **27**, 121–129 (1976).
- Carson, S., Bon, S., Vigny, M. & Massoulié, J. *FEBS Lett.* **97**, 348–352 (1979).
- Vigny, M., Di Giambardino, L., Couraud, J. Y., Rieger, F. & Koenig, J. *FEBS Lett.* **69**, 277–280 (1976).
- Koenig, J. & Vigny, M. *C. R. Séanc. Soc. Biol.* (in the press).
- Méndez, B., Inestrosa, N. C. & Lucio, J. V. *IRCS J. med. Sci.* **7**, 168 (1979).
- Hall, Z. W. & Kelly, R. B. *Nature new Biol.* **232**, 62–63 (1971).
- Bon, S., Vigny, M. & Massoulié, J. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Anglister, L. & Silman, I. *J. molec. Biol.* **125**, 293–311 (1978).
- Guth, L. *Physiol. Rev.* **48**, 645–687 (1968).
- Lømo, T. & Slater, C. R. *J. Physiol., Lond.* (in the press).
- Fernández, H. L. & Inestrosa, N. C. *Nature* **262**, 55–56 (1976).
- Younkin, S. G., Brett, R. S., Davey, B. & Younkin, L. H. *Science* **200**, 1292–1295 (1978).
- Lennon, A. M., Vera, C. L., Rex, A. L. & Lucio, J. V. *J. Neurophysiol.* **30**, 1523–1530 (1967).
- Wilson, S. H. et al. *J. Biol. Chem.* **247**, 3159–3169 (1972).
- Inestrosa, N. C., Ramírez, B. U. & Fernández, H. L. *J. Neurochem.* **28**, 941–945 (1977).
- Martin, R. G. & Ames, B. N. *J. biol. Chem.* **236**, 1372–1379 (1961).
- Méndez, B., Inestrosa, N. C. & Lucio, J. V. *Archs Biol. Med. Exp.* **11**, R-197 (1978).
- Betz, W. & Sakman, B. *Nature new Biol.* **232**, 94–95 (1971).
- Betz, W. & Sakman, B. *J. Physiol., Lond.* **230**, 673–688 (1973).
- McMahan, U. J., Sanes, J. R. & Marshall, L. M. *Nature* **271**, 172–174 (1978).
- Taylor, P., Lwebuga-Mukasa, J., Lappi, S. & Berman, H. A. in *Cholinergic Mechanisms and Psychopharmacology* (ed. D. J. Jenden) 239–251 (Plenum, New York, 1978).
- Inestrosa, N. C. & Lucio, J. V. *Archs Biol. Med. Exp.* **11**, R-150 (1978).
- Walker, C. R. & Wilson, B. W. *Nature* **256**, 215–216 (1975).
- Weinberg, C. & Hall, Z. W. *Dev. Biol.* **68**, 631–635 (1979).
- Lømo, T. & Slater, C. R. in *Synaptogenesis* (ed. Tauc, L.) 9–30 (Naturalia and Biologia, Paris, 1976).
- Lucio, J. V. & Vera, C., *Acta physiol. lat. am.* **14**, 289–294 (1964).
- Fernández, H. L., Duell, M. J. & Festoff, B. W. *J. Neurobiol.* (in the press).
- Rosenthal, J. in *Handbook of Physiology* Vol. 1, (eds Brookhard, J. M. & Mountcastle, V. B.) 775–801 (Am. Physiol. Soc., Bethesda, 1977).
- Landmesser, L. *J. Physiol., Lond.* **220**, 243–256 (1972).
- Skau, K. A. & Brimjoin, S. *Nature* **275**, 224–226 (1978).
- Di Giambardino, L. & Couraud, J. Y. *Nature* **271**, 170–172 (1978).
- Fernandez, H. L., Duell, M. J. & Festoff, B. W. *J. Neurochem.* **32**, 581–585 (1979).
- Griffiths, G. W. *J. Cell Sci.* **36**, 361–390 (1979).
- Oh, T. H. *Expl. Neurol.* **50**, 376–386 (1976).
- Oh, T. H. & Markelonis, G. J. *Science* **200**, 337–339 (1978).
- Koenig, J. & Vigny, M. *Nature* **271**, 75–77 (1978).
- Leighton, F. et al. *J. Cell Biol.* **37**, 482–513 (1968).
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).

Aequorin luminescence during activation of single isolated smooth muscle cells

INCREASED cytoplasmic Ca^{2+} supposedly couples excitation to contraction in smooth muscle. But there have been no direct measurements of such an increase and there is little information about the kinetics of changes in intracellular Ca^{2+} during individual contractions of smooth muscle. Thus the basic aspects of excitation-contraction (EC) coupling in smooth muscle are not known with the same degree of certainty that exists for striated muscle¹. For example, there is a long delay between stimulation and contraction in smooth muscle; this delay might result from a slow onset of Ca^{2+} release after a change in surface membrane potential or from the slowness of some other

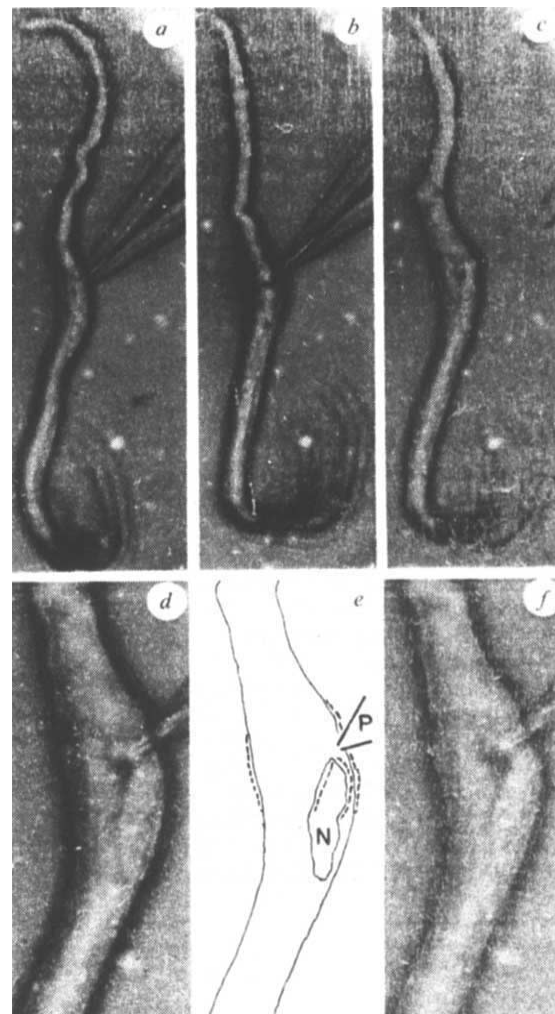
process². If it were possible to correlate measurements of transmembrane potential, intracellular Ca^{2+} and contraction in the same single cell the mechanism of EC coupling in smooth muscle might be better understood. Some progress in this direction has been made by the development of methods for enzymatic isolation of smooth muscle cells and measuring electrical activity and force production or shortening in these single cells^{2–6}. We report here a further approach to the problem by (1) microinjecting the Ca^{2+} indicator aequorin into smooth muscle and (2) directly measuring changes in intracellular Ca^{2+} during EC coupling in single isolated smooth muscle cells. Cytoplasmic Ca^{2+} evidently began to rise without detectable delay following a single electrical stimulus applied to a cell injected with aequorin and reached a peak well before contraction presumably began. The long delay, therefore, between stimulation and contraction seems inherent in the mechanism by which the contractile proteins of smooth muscle are activated.

All experiments were performed on single smooth muscle cells isolated from the stomach muscle of the toad, *Bufo marinus*, by an enzymatic digestion technique that involved the use of trypsin and collagenase. Isolated cells obtained by this technique exhibit physiological, pharmacological, metabolic and ultrastructural properties similar to those of cells in the isolated tissue^{4–6}. But they do not exhibit slow rhythmic spontaneous potential changes once the nerve endings are removed³.

Photomicrographs made during injection of a single smooth muscle cell are shown in Fig. 1. The pipette was left in the cell following the shortening that accompanied insertion. Small volumes of the aequorin solution were then injected by applying gaseous pressure to the butt end of the pipette. The relaxation that resulted when the injected liquid contained EGTA or EDTA was sustained until the cell was stimulated by a current pulse⁷. The swelling that indicated an injection of liquid rapidly subsided when the pressure was released. A cell could typically be injected several times without the swelling itself inducing contraction. The pipette was then withdrawn from a cell, removed from the bath, the quartz-iodine lamp was turned off, and a 25-mm diameter light guide, which extended to the shutter of a photomultiplier⁸, was centred above the cell. Thus, stimulation occurred several minutes after injection whereas aequorin distribution in the cross-section of the injected segment was presumably uniform in only a few seconds (assuming a diffusion coefficient of aequorin of $5 \times 10^{-8} \text{ cm}^2 \text{ s}^{-1}$ (ref. 8); the maximum diameter of 11 cells that were fixed when fully relaxed and subsequently measured by scanning electron microscopy was $4.34 \pm 0.32 \mu\text{m}$). Withdrawal of the pipette did not cause contraction (Fig. 1) and an electrical stimulus could elicit shortening and relaxation after the entire procedure that was not detectably different from a shortening response elicited before the procedure.

The pre-stimulus level of light emission from an aequorin-injected cell was not detectably different from the light recorded during stimulation of a non-injected cell (that is, there seemed to be no 'resting glow' and no change in anode current from the photomultiplier associated with the electrical stimuli themselves). But when a single electrical stimulus was applied to an injected cell a rapid rise in light output was observed. Light output reached a peak within about 10 ms of the onset of stimulation in Fig. 2 for example. The average ($\pm$ s.e.m.) time to peak was 4.8 ± 0.6 ms in 27 of the cells. The response was sometimes large enough to saturate the recording system and the time to peak was not estimated in those cases. Light emission in Fig. 2 initially decayed as a single exponential with a rate constant of 37 s^{-1} but it was still above the pre-stimulus level until at least 250 ms after the onset of the stimulus. Therefore, the rise and most of the fall in luminescence was complete well before mechanical events presumably even began². This is in striking contrast to the behaviour of amphibian skeletal muscle cells where the onset of force development lags no more than a few milliseconds behind the onset of light and light emission declines to only about half its peak level by the time peak force has developed⁸.

Fig. 1 Injection of aequorin into an isolated smooth muscle cell. Photomicrographs made from selected frames of a 16-mm cine film (12 frames per s) taken through an inverted microscope with a Zeiss $\times 40$ phase contrast objective. *Upper row:* Magnification $\times 928$. *a*, The cell before insertion of the micropipette containing aequorin; *b*, 5 s after the insertion of the micropipette; insertion triggered slight shortening of the cell probably caused by a momentary leakage of Ca^{2+} along a shunt resistance in parallel with the input impedance of the cell; *c*, 2 s after withdrawal of the micropipette. Twice between *b* and *c* pressure was applied briefly (~ 1 s) to the micropipette; liquid apparently entered the cell each time since each pressure pulse was accompanied by perceptible but evanescent swelling. The cell remained relaxed following withdrawal but when electrically stimulated it contracted. *Lower row:* Magnification $\times 1,740$. *d*, The same cell after insertion of the micropipette but just before injection of liquid; *e*, the superimposed outlines of the cell before (continuous lines) and during (dashed lines) injection (that is, *d* set on *f*); *f*, during injection of aequorin. The micropipette (P) was always inserted into the uppermost part of the middle of a cell, and when liquid was injected the cine films showed that the most conspicuous distention was usually in a plane perpendicular to the page; furthermore, the nucleus (N) shifted during injection. We also recorded membrane potential responses to electrical stimulation from some isolated cells, and responses ranged from being slow and regenerative during long-duration pulses (see ref. 3) to fast and graded following short-duration pulses (Fig. 2). The entire range of responses could be obtained reproducibly from the same cell and showed that microinjection *per se* did not necessarily abolish the characteristic ability of these cells³ to respond to depolarisation. We also photographed shortening and relaxation of injected cells⁷, which was apparently unchanged and reproducible. This showed that microinjection *per se* did not necessarily change normal EC coupling in these cells. Aequorin was dissolved in either 33 mM EGTA (pH 7.0), 50 mM EDTA (pH 8.0), or quartz distilled water (pH 7.4). Chelating agents can influence aequorin results and their presence may be undesirable as a general rule¹⁷. But EDTA or EGTA in a micropipette prevented or reversed the tendency for a cell sometimes to shorten irreversibly during pipette insertion (Fig. 1). Thus a chelator was often present in our aequorin solutions (for example, the aequorin solutions used to obtain Figs 1 and 3 contained EDTA whereas the solution used to obtain Fig. 2 contained only quartz distilled water). Another advantage to the presence of EDTA or EGTA was that a buffer should oppose a change in $[\text{Ca}^{2+}]$ if it remained undiluted or did not become saturated with Ca^{2+} once inside the cell. Aequorin would then have either detected no change in Ca^{2+} or have indicated a delayed, slow rise as the buffer became saturated. On the contrary, we saw a rapid rise in light emission. Furthermore, the cells contracted and relaxed normally in response to punctate electrical stimulation delivered less than a minute after the injection of EDTA or EGTA which supports the idea that the chelators assume little functional importance after injection⁷. Aequorin solutions were put into filament containing glass micropipettes and pressure injected into individual cells by means similar to those already described for skeletal muscle⁸. But some of the differences in procedure were as follows: (1) the resistances of the micropipettes ranged from 15 to 50 M Ω when filled with 3 M KCl; (2) phase-contrast optics and a quartz-iodine lamp were used for both visual observations and cine-micrography; and (3) cells were stimulated either by a transverse electric field, or by the passage of current between 5 M NaCl in a blunt glass pipette next to a cell and a platinum wire anode positioned at the periphery of the liquid bath. In all the experiments reported here (34 cells) the cells were bathed in a high Ca^{2+} Ringer solution (about $30\times$ normal) that contained (mM) calcium 5.7, potassium 3.0, magnesium 0.98, sodium 45, chloride 162, sulphate 0.98, Tris(hydroxymethyl)amino-methane 2 (pH 7.4), glucose 11, and 6 mg l⁻¹ gentamycin. A high external $[\text{Ca}^{2+}]$ facilitates microelectrode impalements³ and does not change the delay between an electrical stimulus and the onset of contraction². The temperature of the bath ranged from 25° to 34 °C.



The cause for the falling phase of an aequorin response is not certain, but several possibilities can probably be eliminated. First, the decay does not represent total depletion of aequorin since a second stimulus could induce a second flash of light after only one injection; furthermore, the rate constant for the decay of aequorin luminescence from the cells (for example, Fig. 2) was about 20 times higher than the rate constant for decay when aequorin is consumed *in vitro* by rapid mixing with a maximally effective concentration of Ca^{2+} (ref. 9). Second, the rate of quenching of luminescence when aequorin is exposed *in vitro* to Ca^{2+} chelating agents⁹ is about five times higher than the rate constant for decay of light emission from these cells. Thus, the falling phase (Fig. 2) was not determined by the maximum rate imposed by the kinetics of the aequorin reaction. It is also unlikely that the fall of an aequorin response is caused by movement or by a change in the intrinsic optical properties of a cell because the maximum length changes or birefringence changes occur several thousand milliseconds after the maximum change in light emission following electrical stimulation^{5,6}. Calcium binding to sites within the cell and/or the removal of Ca^{2+} from the cytoplasm by membrane associated reuptake mechanisms most likely account for the decay of light¹.

Luminescent responses were not always of the type shown in Fig. 2 (see for example, Fig. 3). This particular atypical response

was much larger and more complex than the type in Fig. 2 and exhibited three phases. Other cells showed only the first phase (I), which could be reproduced by subsequent stimuli¹⁰. The second phase (II) had a rate of decay that was similar to the rate of decay of the response in Fig. 2. The final phase (III) decayed slowly and had a time course consistent with the overall consumption of aequorin⁹. Our hypothesis is that the first and last types of luminescent responses were not only atypical aequorin responses but were associated with atypical membrane responses, that is, responses different from those in Fig. 2 or in ref. 3. Atypical responses might be due to cell injury before or during a stimulus. For example, a cell segment or an entire cell can become inexcitable as a consequence of micropipette penetration or pressure injection yet undergo little or no significant decrease in resting potential¹¹⁻¹⁴. An electrical stimulus, particularly an excessively intense stimulus, could change the membrane potential of such a cell and thus elicit a luminescent response. The total light emitted during responses of the last type (III) were largest and most readily detected but were probably associated with an atypically large rise in intracellular Ca^{2+} . Subsequent stimulation did not elicit further light emission which supports the idea that this type of response (III) is due to overall aequorin consumption. Successive stimuli applied to several such cells in a syncytium could elicit apparently reproducible responses but from different cells.

Observations obtained from single isolated cells preclude this last complexity.

Earlier studies left open the possibility that there might be a long delay between stimulation and the onset of intracellular Ca^{2+} release in smooth muscle². But the rapid rises in light reported here suggest that the delay in EC coupling is not likely to be due to this possibility. Nor is it likely that the delay can be ascribed to the time for diffusion of Ca^{2+} through the cytoplasm. The small diameter of these cells should permit radial gradients in the intracellular $[\text{Ca}^{2+}]$ to be dissipated within 3 to 4 ms after the onset of a stimulus if diffusion is the major factor determining Ca^{2+} distribution. Calcium may carry the current producing a change in membrane potential in these cells³. Therefore, Ca^{2+} entry could account for part if not most of the aequorin signal. The source of the Ca^{2+} cannot be inferred from the present results alone. But one can now discount the possibility that the latency for contraction in smooth muscle results from a delay in the exposure of the contractile filaments to Ca^{2+} . Thus the rate limiting step seems to be imposed by a slow reaction during which the contractile proteins in smooth muscle are activated. Others have suggested that a sequence of Ca^{2+} -sensitive phosphorylation reactions are required for the activation of smooth muscle^{15,16}, and our results support the idea that these reactions impose the characteristically long delay in the activation of smooth muscle contraction.

We need more information to infer the sites responsible for both the rise and fall in intracellular Ca^{2+} following stimulation of these cells. But now that one can inject aequorin into smooth muscle cells these and other questions about the mechanism of EC coupling in smooth muscle can perhaps be answered.

This work was supported by the PHS (grants HL 14523 to F.S.F. and N.S. 14268 to S.R.T.) and the American Heart Association (77-983 to S.R.T.). H.H.S. was supported by PHS

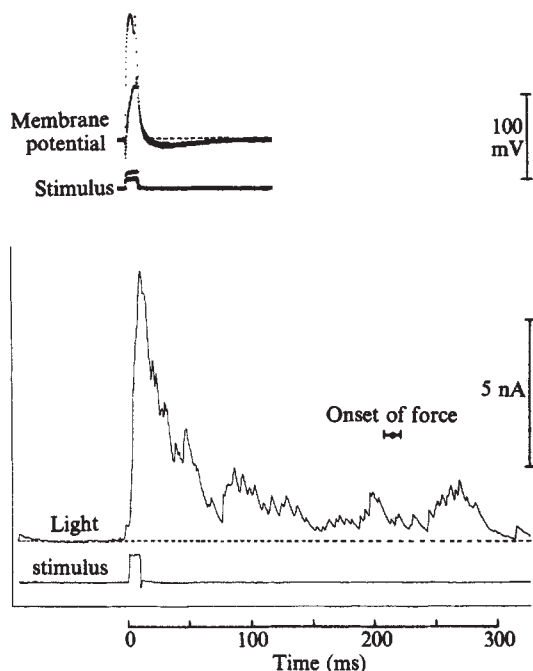


Fig. 2 Light emission and membrane potential change from single, isolated smooth muscle cells stimulated by 10-ms pulses. *Lower traces:* Light emission from a single cell following pressure micro-injection of aequorin and electrical stimulation. The aequorin response was initiated by a pulse applied via wires lying parallel to one another along either side of the cell². The filled circle and bar labelled onset of force indicates the average delay ($\pm$ s.e.m.) between a single electrical pulse and the first detectable sign of mechanical activity². *Upper traces:* Membrane potentials recorded through the same microelectrode (filled with 3 M KCl) that was used to inject the stimulus current into the cell. Field stimulation, on the other hand, either distorted the membrane response recorded via a microelectrode or saturated the recording amplifier. The stimulus duration, solution composition, temperature, and other experimental conditions were the same as those used to obtain the lower aequorin traces.

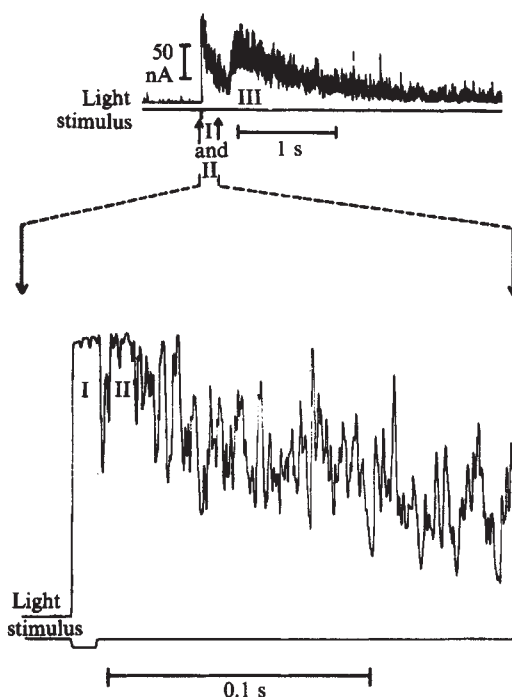


Fig. 3 An unusually large and complex light response from an isolated smooth muscle cell injected with aequorin and then stimulated by a single 10-ms pulse. The entire response is shown on the slower time scale (upper traces). The initial part of the same response is shown on a faster time scale in the lower traces. The initial rise and fall in light (I) is associated with the period of the stimulus. The second rise and fall in light (II) has a time course similar to the response shown in Fig. 2. The third rise and fall in light (III) is the slowest of the three, and has a time course consistent with the partial or total consumption of the aequorin in the cell⁹.

NS 05840. This work was done during the tenure of a Research Career Development Award to F.S.F. (K4HL 00048) and an Established Investigatorship of the American Heart Association to SRT (74-188). F.S.F. was also supported by the Muscular Dystrophy Association. We thank F. G. Prendergast, J. J. Singer and J. V. Walsh for discussions.

F. S. FAY*
H. H. SHLEVIN
W. C. GRANGER JR
S. R. TAYLOR

Department of Pharmacology,
Mayo Foundation and Graduate School of Medicine,
Rochester, Minnesota, 55901

Received 18 December 1978; accepted 25 June 1979.

* Permanent address: Department of Physiology, University of Massachusetts Medical School, Worcester, Massachusetts 01605.

- Somlyo, A. P., Somlyo, A. V., Shuman, H., Sloane, B. & Scarpa, A. *Ann. N.Y. Acad. Sci.* **307**, 523-542 (1978).
- Fay, F. S. *Nature* **265**, 553-556 (1977).
- Singer, J. J. & Walsh, J. V. Jr in *Excitation-Contraction Coupling in Smooth Muscle* (eds Casteels, et al.) 53-60 (Elsevier, Amsterdam, 1977).
- Fay, F. S. *Cold Spring Har. Conf. Cell Motility* **3**, 185-201 (1976).
- Bagby, R. M. *Am. J. Physiol.* **227**, 789-793 (1974).
- Fisher, B. A. & Bagby, R. M. *Am. J. Physiol.* **232**, C5-C14 (1977).
- Fay, F. S. & Taylor, S. R. *Biophys. J.* **21**, 184a (1978).
- Blinks, J. R., Rüdel, R. & Taylor, S. R. *J. Physiol., Lond.* **277**, 291-323 (1978).
- Hastings, J. W., Mitchell, G., Mattingly, P. H., Blinks, J. R. & van Leeuwen, M. *Nature* **222**, 1047-1050 (1969).
- Granger, W. C., Shlevin, H. H., Fay, F. S. & Taylor, S. R. *Fedn Proc.* **38**, 972 (1979).
- Baker, P. F., Hodgkin, A. L. & Shaw, T. I. *J. Physiol., Lond.* **164**, 330-354 (1962).
- Hammer, M. G. & Sheridan, J. D. *J. Physiol., Lond.* **275**, 495-505 (1978).
- Kater, S. B., Nicholson, C. & Davis, W. J. in *Intracellular Staining in Neurobiology* (eds Kater, S. B. & Nicholson, C.) 307-325 (Springer, New York, 1973).
- Muller, K. J. & McMahan, U. J. *Proc. R. Soc. B* **194**, 481-499 (1976).
- Aksoy, M. O., Williams, D., Sharkey, G. M. & Hartshorne, D. J. *Biochem. biophys. Res. Commun.* **69**, 35-41 (1976).
- Sobieszek, A. *Eur. J. Biochem.* **73**, 477-483 (1977).
- Blinks, J. R., Prendergast, F. G. & Allen, D. G. *Pharmac. Rev.* **28**, 1-93 (1976).

Light-activated drug confirms a mechanism of ion channel blockade

A VARIETY of mechanisms underlie the pharmacological blockade of membrane excitability. Some drugs seem to reduce the frequency at which ion channels open; a good example is the effect of curare on acetylcholine receptor channels at normal resting potentials. Another sort of mechanism may account for the action of many local anaesthetics and related drugs containing charged ammonium groups. It is postulated that such molecules block transmembrane currents as they bind to sites within open ion channels, much like a plug in a drain, with the important difference that the events occur on a millisecond time scale. This model, which we shall call 'open-channel blockade', was first applied to the effect of internal tetraethylammonium ions on K^+ channels in squid axon¹ and more recently to similar actions of local anaesthetics on acetylcholine receptor channels²⁻⁴ and on electrically excitable Na^+ channels⁵. (Curare seems to exert an additional open-channel blockade at high negative potentials^{6,7}.) The concept of open-channel blockade would receive direct experimental support from the demonstration that the blockade is exerted even if the blocking molecule is not bound to the channel (or indeed is not present at all) until after the channel opens. Such a demonstration is made possible by a drug that (1) blocks acetylcholine receptor channels in *Electrophorus* electroplaques and (2) is created, in less than a millisecond, by a flash of light.

Bieth *et al.*⁸ have described the photoisomerisable compound, *N*-*p*-phenylazophenylcarbamylcholine iodide (EW-1, Fig. 1). The *cis* isomer can be formed from the *trans* isomer by exposure to light of wavelength 300–500 nm and spontaneously re-isomerises with a time constant of approximately 120 s at room temperature and neutral pH. A solution can be maintained in a predominantly *cis* state by steady exposure to 360-nm light. When *cis*-EW-1 is applied to the innervated face of a voltage-clamped *Electrophorus* electroplaque, neurally evoked postsynaptic currents (PSCs)⁹⁻¹¹ decay more rapidly than normal,

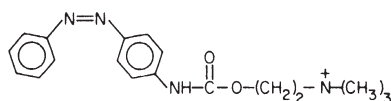


Fig. 1 *N*-*p*-phenylazophenylcarbamylcholine (EW-1), a light-activated cholinergic blocking drug.

the effect reaching fourfold at a concentration of 50 μ M. *cis*-EW-1 also affects the response to bath-applied agonist, causing diminution of agonist-induced currents at equilibrium and rapid components in voltage-jump relaxations. For both PSCs and equilibrium agonist-induced currents, the blockade is more pronounced at more negative membrane potentials. Such effects typify the blockade of acetylcholine receptors by local anaesthetics and other suspected open-channel blockers¹²⁻¹⁷. Similar, but much weaker, effects are seen with *trans*-EW-1.

The light-flash experiments made use of the xenon flash tube and the optical system described by Nass *et al.*¹⁸; the flash was 90% complete within 0.5 ms (individual molecules probably isomerise within less than a microsecond after absorbing a photon). The preparation was bathed in *trans*-EW-1 and kept in the dark between flashes. In experiments on neurally evoked PSCs, a flash accelerated the decay phase of the PSC (Fig. 2). Very few channels open during the decay phase¹⁹⁻²²; therefore the newly created *cis*-EW-1 molecules blocked channels that were already open. After the flash, the PSC decayed with at least two exponential components. We have quantitative data for only the fastest component, which accounted for more than half the decay amplitude in all conditions. This component's rate

constant increases at higher flash intensities, suggesting that it arises from binding of newly created *cis*-EW-1 molecules. Furthermore, the increase is linearly proportional to the number of photoisomerisations (Fig. 3), as expected if each channel can be blocked by the binding of just one *cis*-EW-1 molecule.

Steady conductances are induced by bath-applied agonist. If EW-1 was present as well in the bathing solution, a flash produced a relaxation to a smaller conductance (Fig. 4). As in the PSC experiments, the light-flash relaxation had at least two exponential components. The amplitude of the fastest component was graded with the flash intensity and its rate

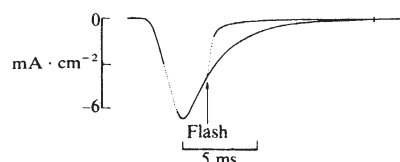


Fig. 2 A light-flash experiment with neurally evoked PSCs. Superimposed traces from two successive voltage-clamp episodes, at an interval of 0.3 s, in 10 μ M *trans*-EW-1. During the second episode, a flash was delivered, isomerising at least 80% of the *trans*-EW-1 molecules to the *cis* isomer. Voltage -150 mV. Temperature 16 °C. Cell 57-42-T34.

constant was greater than that of a voltage-jump relaxation just before the flash, again showing that the blockade arises from an increased rate of channel termination rather than from a decreased rate of channel opening. As a percentage decrease of the conductance, the fast component had the relative size 1.0:0.6:0.35 for the agonists suberyldicholine, acetylcholine, and carbacol. (The experiments were carried out after pretreatment with methanesulphonyl fluoride to block acetylcholinesterase¹⁰.) Thus, the initial component of the blockade is greater for agonists with a greater channel duration^{10,23}, as expected if the newly created *cis*-EW-1 molecules suddenly increase the rate constant for channel closing rate by a fixed amount (in ms^{-1}), with no dependence on the nature of the bound agonist.

The light-flash relaxations had a spectral dependence similar to that of the *trans* $\rightarrow$ *cis* photomerisation of EW-1. When the experiments were carried out using *cis*-EW-1 (rather than *trans*-EW-1), flashes had little additional effect. The signals are specific to acetylcholine receptor channels: light flashes have no effect on electrically excitable Na currents, even in the presence of EW-1.

There are interesting complications in these experiments. (1) The flash-induced blockade in a given *trans*-EW-1 solution is

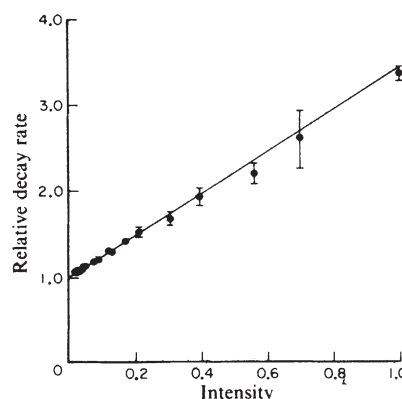


Fig. 3 PSC kinetics as a function of flash intensity. Data from trials like that of Fig. 3, in 5 μ M EW-1. Vertical axis, rate constant of fastest PSC component after the flash, normalised to the value for the immediately preceding trial without a flash. Flash intensity was varied with neutral-density filters; intensity of 1.0 corresponds to 40% *trans* $\rightarrow$ *cis* photoisomerisation. Error bars show $\pm$ s.d. Temperature 10 °C. Cell 57-61, 62.

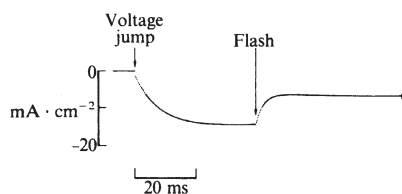


Fig. 4 Voltage-jump and light-flash relaxations in the presence of acetylcholine and 5 μ M EW-1. Agonist-induced current is shown; passive and capacitive currents have been subtracted. Seven milliseconds after the start of the episode, the voltage was jumped from +50 to -150 mV. At about the midpoint of the episode, a flash was delivered, producing a conductance decrease. Flash intensity was the same as for the experiment of Fig. 2. Temperature 10 °C. Cell 61-72, T70, 71.

much greater than the steady-state blockade in a nearly pure *cis*-EW-1 solution of equal concentration. (2) Most of the blockade is transient, disappearing with a time constant of about 300 ms after the flash, even though *cis*-EW-1 has a lifetime of at least 40 s over the pH range from 2-13. These facts suggest that *trans*-EW-1 accumulates in the synaptic cleft; the flash would transiently produce a local *cis*-EW-1 concentration in excess of that in the bathing solution. The site of accumulation is probably not the catalytic site of acetylcholinesterase⁸, because in experiments with carbachol, the light-flash relaxations did not change significantly after the enzyme was inactivated with diisopropyl fluorophosphate⁹ or with methanesulphonyl fluoride¹⁰. The site of accumulation could be the peripheral anionic site of acetylcholinesterase²⁴, the acetylcholine receptors themselves, or some other component of the membrane or synaptic cleft. We do not know whether *cis*-EW-1 accumulates at the same site.

We have obtained very similar results using *N*-*p*-phenylazophenyl-*N*-methyl carbamylcholine and *N*-*p*-phenylazophenyl-*N*-phenyl carbamylcholine²⁵. Like other experiments with light-activated compounds, these studies extend our knowledge of drug-receptor mechanisms^{18,25,29}. With similar techniques it may be possible to test another interesting facet of the 'open-channel blockade' model, namely that the channel conducts again as the blocking drug leaves its binding site^{4,16}.

We thank D. Williams for performing the dissections, M. Walsh for help with the equipment, and J. Nerbonne for discussion. Supported by the NIH (RCDA NS-272 to H.A.L. and grant NS-11756), by the Muscular Dystrophy Association (post-doctoral fellowship to M.M.N. and grant-in-aid), and by the NSF (grant PCM 74-02140).

HENRY A. LESTER
MAURI E. KROUSE
MENASCHE M. NASS

Division of Biology,
California Institute of Technology,
Pasadena, California 91125

NORBERT H. WASSERMANN
BERNARD F. ERLANGER

Department of Microbiology,
Columbia University,
Cancer Center/Institute for Cancer Research,
College of Physicians and Surgeons,
New York, New York 10032

Received 5 March; accepted 6 June 1979.

1. Armstrong, C. M. *J. gen. Physiol.* **50**, 491-503 (1966).
2. Steinbach, A. B. *J. gen. Physiol.* **52**, 144-161 (1968).
3. Steinbach, A. B. *J. gen. Physiol.* **52**, 162-180 (1968).
4. Adams, P. R. *J. Physiol., Lond.* **246**, 61-63P (1975).
5. Strichartz, G. R. *J. gen. Physiol.* **62**, 37-57 (1973).
6. Katz, B. & Miledi, R. *Proc. R. Soc. B203*, 119-133 (1978).
7. Colquhoun, D., Dreyer, F. & Sheridan, R. E. *J. Physiol., Lond.* (in the press).
8. Bieth, J., Wassermann, D., Vratsanos, S. M. & Erlanger, B. F. *Proc. natn. Acad. Sci. U.S.A.* **66**, 850-854 (1970).
9. Sheridan, R. E. & Lester, H. A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3496-3500 (1975).
10. Sheridan, R. E. & Lester, H. A. *J. gen. Physiol.* **70**, 187-219 (1977).
11. Lester, H. A., Koblin, D. D. & Sheridan, R. E. *Biophys. J.* **21**, 181-194 (1978).

12. Furukawa, T. *Jap. J. Physiol.* **7**, 199-212 (1957).
13. Lester, H. A., Changeux, J.-P. & Sheridan, R. E. *J. gen. Physiol.* **65**, 797-816 (1975).
14. Beam, K. G. *J. Physiol., Lond.* **258**, 279-300 (1976).
15. Adams, P. R. *J. Physiol., Lond.* **268**, 291-318 (1977).
16. Neher, E. & Steinbach, J. H. *J. Physiol., Lond.* **277**, 153-176 (1978).
17. Koblin, D. D. & Lester, H. A. *Molec. Pharmac.* (in the press).
18. Nass, M. M., Lester, H. A. & Krouse, M. E. *Biophys. J.* **24**, 135-160 (1978).
19. Magleby, K. L. & Stevens, C. F. *J. Physiol., Lond.* **223**, 151-171 (1972).
20. Magleby, K. L. & Stevens, C. F. *J. Physiol., Lond.* **223**, 173-197 (1972).
21. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655-691 (1973).
22. Wathey, J. C., Nass, M. M. & Lester, H. A. *Biophys. J.* **27**, 145-164 (1979).
23. Katz, B. & Miledi, R. *J. Physiol., Lond.* **230**, 707-717 (1973).
24. Rosenberry, T. in *Adv. Enzymol.* **43**, 103-218 (1975).
25. Deal, W. J., Erlanger, B. F. & Nachmansohn, D. *Proc. natn. Acad. Sci. U.S.A.* **64**, 1230-1234 (1969).
26. Bartels, E., Wassermann, N. H. & Erlanger, B. F. *Proc. natn. Acad. Sci. U.S.A.* **68**, 1820-1823 (1971).
27. Erlanger, B. F. *A. Rev. Biochem.* **45**, 267-283 (1976).
28. Lester, H. A. & Chang, H. W. *Nature* **266**, 373-374 (1977).
29. Wassermann, N. H., Bartels, E. & Erlanger, B. F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 256-259 (1979).

Occlusion of K ions in the unphosphorylated sodium pump

POST, Hegyvary and Kume¹ found that the rate of rephosphorylation of (Na⁺ + K⁺)ATPase that had just been dephosphorylated depended on which congener of K (Rb or Li) had been used to catalyse the hydrolysis of the phosphoenzyme, even when the experiments were done in such a way that the composition of the medium during the rephosphorylation was constant. To explain the ability of the enzyme to 'remember' which kind of ion had catalysed the hydrolysis, they postulated that during the hydrolysis the catalysing ions became occluded within the enzyme, and were released only later after a slow conformational change. Furthermore, as ATP in high concentrations accelerated the rephosphorylation, they supposed that the slow conformational change was accelerated by the binding of ATP to a low-affinity site on the enzyme. Karlsh, Yates and I.M.G.² suggested that the hypothetical 'occluded-K form' of Post *et al.* was identical with the form of unphosphorylated (Na⁺ + K⁺)ATPase that predominates in K-containing Na-free media; they found that the conversion of that form (E₂K) into the form (E₁Na) that predominates in high-Na media was very slow ($k \approx 0.25 \text{ s}^{-1}$ at 20 °C), and was accelerated by nucleotides acting with a low affinity and without phosphorylation. Karlsh *et al.* also studied the conformational change in the reverse direction—E₁Na → E₂K—and concluded that K ions combine with the E₁ form of the enzyme at sites with a relatively low affinity. They therefore suggested that the occluded-K form could be generated either following the binding of K⁺ ions at low-affinity sites, probably on the intracellular surface of the membrane, or by the hydrolysis of phosphoenzyme catalysed by K ions bound to high-affinity sites on the extracellular surface of the membrane^{2,3} (see Fig. 1). They pointed out that, if their hypothesis were correct, it could account for inward and outward movements of K ions through the sodium pump, and could explain why orthophosphate and ATP (or its non-phosphorylating analogues) are required for K-K exchange. However, the similarities between the properties of the occluded-K form and the form that predominates in K-containing, Na-free media could be coincidental. Furthermore, the idea of an occluded-K form is also open to doubt, as the experiments of Post *et al.* prove only that the product of the hydrolysis of phosphoenzyme differs depending on whether Rb or Li ions catalyse the hydrolysis, not that the ions become trapped inside. Both problems could be resolved if it were possible to show that, in Na-free media, K ions or Rb ions binding to the unphosphorylated form of the enzyme were indeed occluded. We report here experiments showing that they are.

Briefly, the procedure followed in these experiments was (1) to suspend a partially purified suspension of (Na⁺ + K⁺)ATPase in a medium free of Na and Mg and containing 100 μ M ⁸⁶RbCl, with or without nucleotide; (2) to pass the suspension rapidly down a cation-exchange column loaded with Na; (3) to measure

Table 1 Retention of Rb by $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ after the removal of free Rb by passage of the enzyme through a cation-exchange resin

	Nucleotide in suspension medium	Protein content of effluent (mg)	Rb content of effluent (nmol)	Extra Rb in absence of nucleotide (nmol)	Extra Rb (nmol per mg protein)
Expt 1	None	0.628 ± 0.025	0.992 ± 0.051	0.699 ± 0.051	1.12 ± 0.09
	2 mM ATP	0.657 ± 0.027	0.293 ± 0.003	—	—
Expt 2	None	0.568 ± 0.019	1.160 ± 0.042	0.825* ± 0.044	1.45 ± 0.09
	2 mM ATP	0.572 ± 0.018	0.351 ± 0.025	—	—
	2 mM ADP	0.582 ± 0.010	0.318 ± 0.012	—	—

One-mg quantities of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ (specific activity 12 μmol per mg per min), suspended in 0.5-ml portions of a medium containing 100 mM Tris/Tris Cl (pH 7.4 at 20 °C), 15 mM histidine, 1 mM EDTA (Tris salt) and 100 μM $^{86}\text{RbCl}$, with or without nucleotide, were passed through columns of Biorad Biorex 70 carboxylic resin (50–100 mesh) in the Na form. Each column was prepared by putting the resin, suspended in 100 mM HEPES (Na salt, pH 7.4), in the barrel of a 1-ml disposable tuberculin syringe (cross-sectional area 0.17 cm^2) whose tip was loosely plugged with glass fibre and fitted with a nylon tap. The final bed volume was 0.5 ml, and the total capacity was 1.65 mEq. After the resin had packed down, the tap was opened until the surface of the Na HEPES solution reached the top of the resin. The upper part of the syringe was then filled with water, and acid-washed sand was allowed to settle through the water to form a layer 1 mm thick on top of the resin. This was to prevent contamination of the enzyme suspension with Na from the resin. The water above the sand was removed by aspiration, and the upper part of the syringe was washed twice with water. The enzyme suspension was slowly run into the syringe without disturbing the sand layer, and the syringe was mounted vertically on the moveable platform of a Palmer stand, whose coarse screw (pitch 5 mm) could be turned at predetermined rates by an electric motor acting through a friction clutch. The piston was inserted into the barrel so that about 0.25 ml of air was trapped above the enzyme suspension; a small hole drilled through the wall of the barrel, at a point corresponding to 1.3 ml on the (extrapolated) scale, allowed this to be done without any significant rise in pressure. The tap was opened and the motor switched on. Suitable stops had been arranged so that, as the syringe was raised by the motor, the piston was pushed in, the movement beginning only after the motor had reached its set speed, and terminating abruptly just before the piston reached the surface of the resin. The flow rate was 0.22 ml s^{-1} and the temperature 21 °C. The effluents were analysed for total protein* and for ^{86}Rb . The figures in the second and third columns of the table are each the mean of three determinations $\pm$ s.e.m.

* Calculated by subtracting the average of the Rb contents of the ATP and ADP samples from the average Rb content of the nucleotide-free samples.

the amounts of radioactivity and of protein emerging from the column. Rb was used rather than K, partly because ^{86}Rb has a more convenient half life than ^{42}K , and partly because the data of Post *et al.*¹ suggest that Rb is released from the occluded form slightly more slowly than K.

Details of the procedure are given in Table 1, but it is convenient to consider some of them here. The main difficulty was to find conditions which allowed the column to remove free Rb from the suspension completely enough for the small amount of Rb occluded within the enzyme not to be masked, and quickly enough for the enzyme to emerge before too much of the occluded Rb had been released. $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ from pig kidney outer medulla was used because it can be prepared conveniently about 30% pure in reasonably large quantities⁴, and the purified preparation is not vesiculate. When the suspending media were to contain ATP, the ATP was added only 1 or 2 min before the experiment, so that hydrolysis (in the absence of Mg and Na, and in the presence of 1 mM EDTA) should have been negligible. With 100 μM Rb, 2 mM ATP or

ADP is more than sufficient to displace the equilibrium between E_2 and E_1 forms of the enzyme virtually completely in favour of the E_1 form⁵; in the presence of nucleotide, any occluded Rb would therefore have been released before the enzyme was added to the column.

Preliminary experiments showed that, with the conditions chosen: (1) in the absence of enzyme, the column removed more than 99% of ^{86}Rb from 0.5 ml of a solution containing 100 μM $^{86}\text{RbCl}$, 100 mM Tris/Tris Cl (pH 7.4) and 1 mM EDTA (Tris salt); the small fraction of Rb remaining in the effluent was not significantly affected if the solution also contained 2 mM ATP; (2) when 1 mg of enzyme was added to the solution, $60 \pm 4\%$ of the enzyme was recovered in the effluent; the presence of 2 mM ATP did not significantly affect the recovery of enzyme.

The results of two separate experiments, each carried out in triplicate, are summarised in Table 1. In the presence of nucleotides the amounts of Rb that emerged from the columns were about 0.6% of the total, and not significantly different from the amounts of Rb that emerged from the columns in control experiments without enzyme. In the absence of nucleotides, an extra amount of Rb was carried through the columns, which is just what we should expect if the enzyme contained occluded

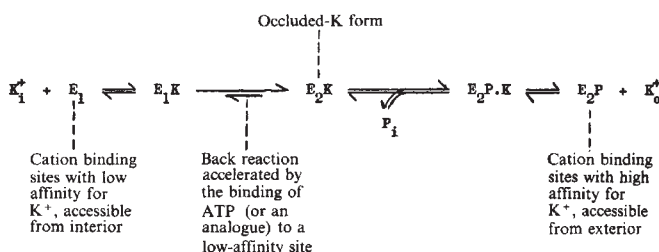


Fig. 1 Scheme showing the hypothesis of Karlish, Yates and I.M.G.², in which two routes to the occluded-K form are coupled back-to-back to produce a chain of reactions capable of carrying K ions through the membrane. Note: (1) E_2P in this scheme is identical with E_2P in the Albers-Post scheme (see Fig. 1 of ref. 8); E_2K in this scheme is identical with E_2K in Fig. 4 of ref. 2, and also with $\text{E}_2(\text{K})$ in Fig. 5 of ref. 2 and Fig. 1 of ref. 8. (2) In the present experiments we are concerned only with the unphosphorylated forms of the enzyme. (When ATP was present, phosphorylation was prevented by the lack of Mg and, for most of the time, also of Na.)

Rb. Before we accept that explanation, however, we must see how this extra amount compares with the amount of occluded Rb that the enzyme might be expected to carry through the columns.

In the last column of Table 1, the extra Rb that appeared in the effluents in the absence of nucleotide has been expressed as nmol Rb per mg protein in each effluent. The specific activity of the enzyme used in these experiments was 12 μmol per mg protein per min, at 37 °C in standard conditions, and as enzyme approaching 100% purity has a specific activity of 40 μmol per mg protein per min in identical conditions, the purity of our enzyme must have been about 30%. If the molecular weight of the enzyme is assumed to be 250,000, our preparation must have contained 1.2 nmol of enzyme per mg protein. If each enzyme molecule can occlude two Rb ions, the maximum amount of occluded Rb would therefore be 2.4 nmol per mg protein. The amount of extra Rb to be expected in the column effluents is, however, less than this for two reasons: first, at an Rb concentration of 100 μM , not all the enzyme would have been in the E_2 form initially, and second, some of the occluded Rb would have been released during the time the enzyme was passing down the column. The relevant data for Rb are not available, but if we assume that Rb behaves like K, the fraction of enzyme in the E_2 form initially would have been about two-thirds⁶. The rate constant for the conversion of E_2K to E_1K , which is thought to be the rate-limiting step for the release of occluded K, has

been estimated to be about 0.25 s^{-1} at 20°C in the absence of nucleotides^{2,4}. Using the data in Table 1 of ref. 1, we calculate that the rate constant for the conversion of E_2Rb to E_1Rb is probably about 0.64 times as great. The amount of extra Rb to be expected per mg protein in the effluent is therefore $2.4 \times 0.67 \times \exp(-0.25 \times 0.64t) \text{ nmol}$, where t is the average time that enzyme in the effluent has spent passing through the resin. We cannot give a precise value for t , but it was almost certainly between 0.8 and 2.0 s. The expected amount of extra Rb is therefore between 1.17 and 1.41 nmol per mg protein. Given the uncertainties in the calculation, and the experimental errors, the agreement between the predicted and the observed values is remarkably good. We therefore conclude that the extra Rb carried through the columns with the enzyme in the absence of nucleotides does represent Rb occluded within the E_2 form of the enzyme.

We are now trying to adapt the rapid ion-exchange technique used in the above experiments to test the hypothesis^{2,3} that Na ions are occluded within the phosphorylated ($\text{Na}^+ + \text{K}^+$)ATPase when it is in the E_1P conformation, the conformation believed to be the immediate product of phosphorylation of the protein by ATP.

We thank Mr R. H. Cook and Mr W. Smith for technical assistance and the MRC for financial support.

L. A. BEAUGÉ
I. M. GLYNN

Physiological Laboratory,
University of Cambridge,
Cambridge, UK

Received 3 May; accepted 5 July.

- Post, R. L., Hegyvary, C. & Kume, S. *J. biol. Chem.* **247**, 6530–6540 (1972).
- Karlish, S. J. D., Yates, D. W. & Glynn, I. M. *Biochim. biophys. Acta* **525**, 252–264 (1978).
- Glynn, I. M., Karlish, S. J. D. & Yates, D. W. in *2nd Int. Conf. Properties and Functions of Na,K-ATPase* (ed. Skou, J. C.) (Academic, New York, in the press).
- Jørgensen, P. L. *Biochim. biophys. Acta* **256**, 36–52 (1974).
- Beaugé, L. A. & Glynn, I. M. *J. Physiol. Lond.* (in the press).
- Karlish, S. J. D. & Yates, D. W. *Biochim. biophys. Acta* **527**, 115–130 (1978).
- Lowry, O. H., Rosebrough, N. U., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
- Karlish, S. J. D., Yates, D. W. & Glynn, I. M. *Biochim. biophys. Acta* **525**, 230–251 (1978).

Highly active cyclic and bicyclic somatostatin analogues of reduced ring size

THE biological activity of an organic compound represents the sum of several properties, including solubility, absorption, transport, plasma protein binding, metabolism and receptor binding. The degree of molecular flexibility may affect these properties in different ways. Our approach to the design of somatostatin analogues with reduced susceptibility to metabolic inactivation has been both to eliminate those amino acids which are not required for biological activity and to increase the rigidity of the molecule. We report here the preparation of conformationally constrained analogues of somatostatin (Fig. 1, IIa and III) which are highly active inhibitors of the release of insulin, glucagon and growth hormone *in vivo*. Analogue III (Fig. 1), which retains only four of the amino acids of the natural hormone (sequence 7–10), is relatively resistant to the action of trypsin *in vitro*.

We have already reported¹ the preparation of analogue I, a highly active bicyclic analogue of somatostatin that has a covalent bridge between the β -carbons of amino acids 5 and 12 (numbers assigned to amino acid residues of the analogues correspond to those of somatostatin). On the basis of the high biological activity of analogue I, we concluded that the corresponding β -carbons in somatostatin are in close proximity when somatostatin is bound to the receptors involved in the

Table 1 Characterisation data for analogues

Amino acid analysis				[α] _D ²⁵ (conc.) [†]	R_F ‡ (solvent system)§
IIa	Lys 1.00	Thr 0.98	Phe 3.05	−49.2 (0.15)	0.79 (A), 0.32 (B)
	Trp 0.97	Aha 0.99*			
IIb	Lys 1.00	Thr 0.70	Phe 0.98	−35.5 (0.2)	0.58 (A), 0.13 (B)
	Trp 1.00¶	Aha 1.01			
	Cys* 1.78				
IIc	Lys 1.00	Thr 0.96	Phe 3.04	−43.2 (0.1)	0.79 (A), 0.35 (B)
	Trp 0.97	Aha 1.00*			
III	Lys 0.99#	Thr 1.01	Phe 1.00	−64.7 (0.25)	0.69 (A), 0.23 (B)
	Trp 1.00¶	Aha 1.03	Cys 1.89		

* Methanesulphonic acid-catalysed hydrolysis for 20 h at 110°C . † In 50% aqueous AcOH. ‡ Silica gel, Analtech. § (A) 10:5:1:3, AcEtO/pyridine/AcOH/H₂O; (B) 80:20:2, CHCl₃/MeOH/conc. NH₄OH. || When Thr (or Ser) precede cysteine in a peptide we have observed simultaneous destruction of both amino acids. Racemisation of Ile in the sequence Ile-Cys has been reported to involve thiazoline formation¹⁰. Presumably, thiazoline formation would allow further reactions when a β -hydroxyl, capable of elimination, is present. ¶ Determined by UV. # after 20 hours 6M HCl hydrolysis at 110°C .

inhibition of the release of insulin, glucagon and growth hormone. Reduction of the disulphide bridge to give the monocyclic, dihydro derivative did not explain the high activity of analogue I *in vitro* and *in vivo*¹. At that time our attention was focused on amino acids 7–10 of somatostatin. These are important in the interaction with receptors whereas amino acids 5, 6, 11 and 12 are not. It had been reported earlier that amino acids 4 and 13 could be deleted individually while retaining biological activity². These observations suggested that the amino acids to the left of the disulphide bridge of analogue I could be deleted and the disulphide bridge replaced by methylene groups while retaining biological activity. We synthesised analogue IIa, cyclo-(Aha-Phe-Phe-D-Trp-Lys-Thr-Phe), also designed as an analogue of I incapable of *in vivo* reduction^{3,4} (Aha is ω -aminoheptanoic acid). A disulphide-containing analogue of ring size similar to that of analogue IIa has been reported to have significant biological activity⁵.

The minimum requirements for biological activity can be tested further by replacing phenylalanines 6 and 11 of analogue IIa by cystine, as we have reported for an analogue with the full ring size of somatostatin¹. Therefore, we also synthesised the bicyclic analogue III, cyclo-(Aha-Cys-Phe-D-Trp-Lys-Thr-Cys), which retains only amino acids 7–10 of somatostatin.

These analogues were prepared as outlined in Fig. 2. The L-Trp analogue of IIa, cyclo-(Aha-Phe-Phe-Trp-Lys-Thr-Phe) (IIc), was prepared in a similar way. Analytical results and physical properties are summarised in Table 1. The NMR spectra at 300 MHz supported the structures by demonstrating

Table 2 Inhibition of release (relative potency of somatostatin = 1)

Compound	Glucagon	Insulin	Growth hormone		Gastric secretion
			<i>In vitro</i>	<i>In vivo</i>	
I	1.4 (0.26–12.40)	1.53 (0.92–2.66)	0.37 (0.29–0.48)	1.88 (0.35–8.33)	0.10
IIa	0.86 (0.44–1.53)	0.88 (0.30–2.45)	0.93 (0.69–1.2)	0.65 (0.20–4.61)	0.03
IIb		0.020 (0.003–0.060)	0.03 (0.02–0.04)	0.14 (0.05–0.29)	0.03
IIc	0.041 (0.004–0.150)	0.10 (0.05–0.19)	0.14 (0.12–0.16)	—	<0.01
III	2.66 (1.32–6.10)	3.50 (2.31–6.38)	1.24 (0.81–1.88)	2.55 (0.99–11.1)	0.05

The biological effects were measured using methods described previously¹. Inhibition of glucagon and insulin release was measured in urethane anaesthetised rats. Growth hormone release inhibition *in vivo* was measured in pentobarbital-treated rats. *In vitro* inhibition of growth hormone release was measured using dispersed pituitary cells. Inhibition of gastric secretion was measured in dogs stimulated by pentagastrin. 95% confidence limits are given in parentheses. For glucagon release by analogue IIb, a low level of activity was observed, with no dose response.

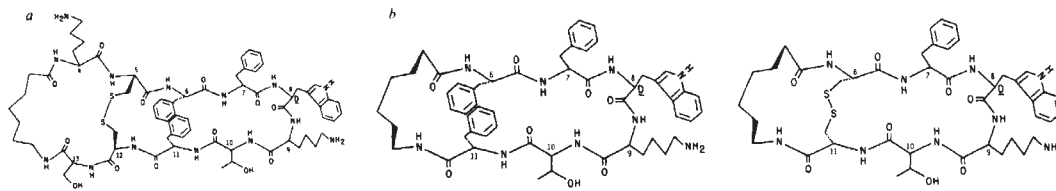


Fig. 1 Structure of analogues I, IIa and III.

normal resonances for all the amino acids, with no unassignable resonances. The only abnormal resonance was the γ -CH₂ of Lys 9, shifted upfield about 1 p.p.m. in analogues IIa, IIb and III, a shift we have attributed to a proximity of the side chains of Trp 8 and Lys 9 in analogues with D-Trp in position 8. The high field resonance in these new analogues lacking Lys 4 permits unequivocal assignment to the CH₂ of Lys 9 as previously proposed⁶.

The biological potencies of the analogues are summarised in Table 2, with the results previously reported for analogue I. In these respects analogues I, IIa and somatostatin are nearly the same except that I and IIa are less potent inhibitors of gastric secretion. It is clear from these and previous studies that inhibition of gastric secretion involves unique structural requirements. The high potency of analogue IIa provides additional evidence that *in vivo* disulphide reduction is not a prerequisite for the biological activity of analogue I.

We have suggested that the roles of Phe 6 and Phe 11 are related, at least in part, to stabilising a peptide conformation through the formation of a hydrophobic bond¹. This conclusion is further strengthened by the high biological activity of analogue III relative to IIb. The disulphide bridge of analogue III seems to be a prerequisite of high biological potency because the lack of bonding between amino acids 6 and 11, either covalent (III) or hydrophobic (IIa), results in lowered potency, as seen in analogue IIb. Once again, the D-Trp 8 analogue IIa is more potent than the L-Trp diastereomer (IIc), as was first observed with somatostatin⁷ and later reported also in the dicarba-series⁸. We have interpreted this higher potency in the D-Trp series as representing a stabilisation of the biologically active conformation⁶. The high activity of analogues IIa and III indicates that the binding of somatostatin to receptors is due chiefly to amino acids 7–10. The potencies reported in Table 2 not only reflect changes in receptor binding but are also a summation of *in vivo* effects, including metabolism. The *in vitro* growth hormone release data give the closest estimation of receptor binding because potential metabolism has been minimised by the use of conditions previously developed for a successful pituitary receptor model⁹.

The bicyclic analogue III was more resistant to enzymatic cleavage, as demonstrated by its relative resistance to trypsin hydrolysis between lysine and threonine *in vitro*. The monocyclic analogue IIa was cleaved by trypsin at about 100 times the rate of hydrolysis of III in identical conditions (1.5×10^{-3} M peptide). The course of hydrolysis was followed using high pressure liquid chromatography (HPLC) to assess the ratio of substrate to product at various intervals. It is interesting that in the presence of 1.5×10^{-3} M mercaptoethanol, analogue III was hydrolysed at about the same rate as IIa. Thus, the disulphide bond adds enzymatic stability apparently through conformational or steric constraint. The lower rate of tryptic hydrolysis was due to reduced binding to trypsin because analogue III did not inhibit tryptic hydrolysis of IIa even when present in a concentration 10-fold higher than that of analogue IIa. The introduction of conformational constraint seems mainly to have increased K_M rather than prevented some needed conformational change along the pathway of hydrolysis.

There was an increased duration of action after subcutaneous administration of analogue III relative to IIa and somatostatin (Fig. 3): the data in Fig. 3 show that 75 min after injection, both IIa and III significantly inhibit the release of growth hormone,

whereas somatostatin does not. At 135 min after injection, only analogue III shows significant inhibition of growth hormone release. Indeed, III significantly inhibited growth hormone release in this system up to 5 h after subcutaneous administration (to be reported in detail elsewhere). Although it seems reasonable that the greater duration of action is a consequence of the enhanced enzymatic stability of analogue III, it may be related in part to solubility differences, giving effectively depot forms of the analogue⁸.

In conclusion, we believe that amino acids 7–10 of somatostatin, when constrained in the correct conformation, contain receptor binding elements sufficient to express the total activity of somatostatin for suppression of the release of insulin, glucagon and growth hormone. Our studies suggest that conformationally constrained structures can be designed to

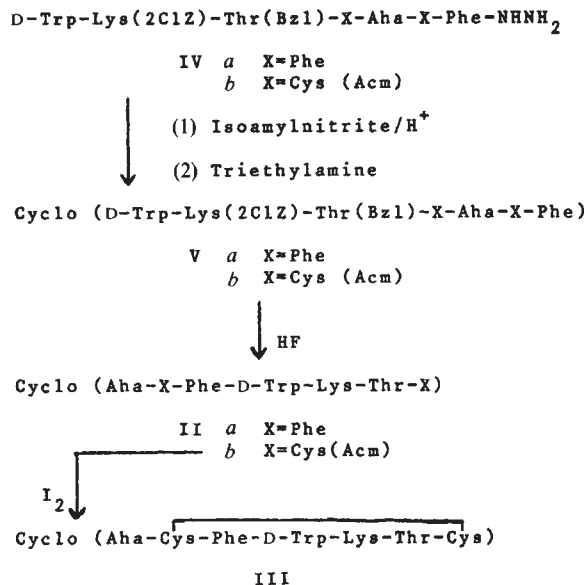


Fig. 2 Route for synthesis of analogues. The experimental details of the synthetic procedure are described in ref. 1 for larger ring monocyclic and bicyclic analogues of somatostatin. IIa was prepared without purification of intermediates. The final product (III) was purified by chromatography on silica 60 using CH₂Cl₂/isopropanol/NH₄OH (concentrated) (60:40:5) as eluant. Isolated yield was 42% of product plus 22% of material showing trace impurities (melting point 234–236). In this study I₂ oxidation of analogue IIb was carried out at a concentration of 3.2 mM in dimethylformamide using 6 equivalents of I₂. Excess I₂ was removed by filtration through silica gel 60, the I₂ being eluted with CHCl₃ and 10% MeOH/CHCl₃ while the peptide was adsorbed. III was eluted with CHCl₃/MeOH/H₂O (6:4:1) and the crude peptide product further purified by filtration through Sephadex G-50F, 50% AcOH, to give an isolated yield of 35%. III was generally prepared without purification of any intermediates. Purified samples of the intermediate IIb were obtained by chromatography on silica gel, eluted with CHCl₃/MeOH/NH₄OH (70:30:3) followed by filtration through Sephadex G-25F (2M AcOH). The properties and analytical data obtained on these synthetic products and cyclo-(Aha-Phe-Phe-Trp-Lys-Thr-Phe) (IIc) are summarised in Table 1. Final products had a purity of $\geq 98\%$ as analysed by TLC and HPLC.

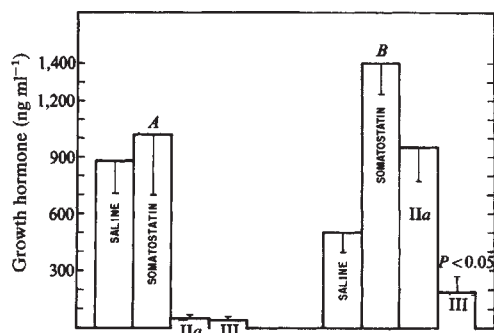


Fig. 3 Duration of action of somatostatin and analogues on the inhibition of growth hormone release. Male Sprague-Dawley rats (190–200 g) were injected with saline or compound subcutaneously. After 1 or 2 h sodium pentobarbital (17 mg per kg intravenously) was injected. The rats were bled 15 min later from the orbital sinus. The plasma was collected and assayed for growth hormone content by radioimmunoassay. The data are presented as the mean $\pm$ s.e.m. (six rats per group). Doses of somatostatin (50 μ g) and analogues IIa (50 μ g) and III (25 μ g) were chosen to attain equipotence. A, 75 min after administration; B, 135 min after administration.

decrease susceptibility to metabolism while retaining biological activity. Because analogue III shows greater stability both *in vivo* and *in vitro*, we believe that the incorporation of conformational constraint is a viable approach to the design of long acting somatostatin analogues.

We thank Mr William J. Paleveda, Jr, Mr Carl Homnick, Mr E. W. Chapin, Miss M. Pastor, and Mrs J. Wittreich for assistance, Dr Byron Arison for NMR spectra, Dr Marylou Torchiana for data on gastric secretion, and Mr Martin Schnall for help with statistical analysis.

DANIEL F. VEBER
FREDERICK W. HOLLY
RUTH F. NUTT
SUSAN J. BERGSTRAND
STEPHEN F. BRADY
RALPH HIRSCHMANN

Merck Sharp and Dohme Research Laboratories,
West Point, Pennsylvania 19486

MONROE S. GLITZER
RICHARD SAPERSTEIN

Merck Institute for Therapeutic Research,
Rahway, New Jersey 07065

Received 2 April; accepted 21 June 1979.

1. Veber, D. F. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 2636–2640 (1979).
2. Sarantakis, D., McKinley, W. A., Jaunakais, I., Clark, D. & Grant, N. H. *Clin. Endocr.* **5**, 275–278S (1976).
3. Rudinger, J. & Jost, K. *Experientia* **20**, 570–571 (1964).
4. Jost, K. & Rudinger, J. *Colln Czech. chem. Commun. Engl. Edn* **32**, 1229–1241 (1967).
5. Vale, W., Rivier, J., Ling, N. & Brown, M. *Metabolism* **27**, Suppl. 1, 1391–1401 (1978).
6. Arison, B. H., Hirschmann, R. & Veber, D. F. *Bioorg. Chem.* **7**, 447–451 (1978).
7. Rivier, J., Brown, M. & Vale, W. *Biochem. biophys. Res. Commun.* **65**, 746–751 (1975).
8. Torchiana, M. L., Cook, P. G., Weise, S. R., Saperstein, R. & Veber, D. F. *Archs int. Pharmac. Ther.* **235**, 170–176 (1978).
9. Schonbrunn, A. & Tashjian, A. H. Jr *J. biol. Chem.* **253**, 6473–6483 (1978).
10. Craig, L. C., Hausmann, W. & Weisiger, J. R. *J. Am. chem. Soc.* **76**, 2839 (1954).

δ -Aminolaevulinic acid is a potent agonist for GABA autoreceptors

ACUTE attacks of the hereditary hepatic porphyrias are frequently accompanied by neuropsychiatric symptoms. The presentation of nervous system involvement varies; in decreasing order of frequency, there may be motor neuropathy, confusion, psychiatric manifestations, hyperexcitability and epileptic-type seizures¹. Psychiatric manifestations include depression, anxiety, insomnia and an organic brain syndrome². Although the aetiology of the neural dysfunction in the acute

attack is unknown, several mechanisms have been suggested at a neurochemical level. It is possible that porphyrin precursors, which are produced in excess by the liver during acute episodes, gain access to the nervous system and exert direct neurotoxic effects^{3–5}; alternatively, a metabolic defect in haem biosynthesis in nerve cells might result in a functional haemoprotein deficiency which becomes critical at the time of the acute attack⁶. The porphyrin precursor, δ -aminolaevulinic acid (δ -ALA), exhibits various effects in mammalian central nervous tissue preparations. It inhibits high-affinity uptake and increases efflux of γ -aminobutyric acid (GABA)³ and L-glutamate⁴ in rat cortical synaptosomes; it also inhibits the (Na^+ + K^+) ATPase isolated from rabbit brain and red blood cells⁷. These effects occur only at relatively high concentrations of δ -ALA (at least 10^{-4} M), and it is very unlikely that such concentrations are reached in the brain during acute porphyric attacks⁸. Recently, there has been strong evidence that release of GABA is subject to negative feedback control through presynaptic receptors on GABAergic terminals^{9–11}. Here we report that δ -ALA inhibits potassium-stimulated release of GABA from preloaded synaptosomes, probably by acting as an agonist at these presynaptic receptors. This effect is evident at concentrations of δ -ALA (10^{-6} M) which are well within levels recorded in the cerebrospinal fluid of porphyric patients showing neuropsychiatric symptomatology¹².

The method used has been published previously¹³. Briefly, fresh synaptosomes prepared from the cerebral cortices of adult male Wistar rats were resuspended in 0.32 M glucose to give a protein concentration of about 5 mg ml⁻¹ and diluted 1:10 with ice-cold incubation medium (final concentrations: 128 mM NaCl; 5 mM KCl; 2.7 mM CaCl₂; 1.2 mM MgSO₄; 0.1 mM aminooxyacetic acid to prevent GABA catabolism; 10 mM Tris-HCl buffer at pH 7.35). One-ml aliquots of the suspension were preincubated at 37 °C for 15 min in a water bath in an air atmosphere. [$2,3$ -³H]GABA of specific activity 54 Ci mmol⁻¹ (Radiochemical Centre) was diluted with unlabelled GABA to give a specific activity of 10 Ci mmol⁻¹ and a small volume (1/100 of the final volume) added to the incubation tubes to a final concentration of 0.5×10^{-6} M. Incubation was continued for a further 10 min after which the synaptosome suspensions were layered on Millipore filters (0.45 μ m pore) resting on filter supports constituting the bottoms of four parallel superfusion chambers¹⁴, and thermostatically maintained at 37 °C. The chambers were connected to a multi-channel peristaltic pump, the excess medium drawn off at maximum flow rate and the filters washed with 10 ml of medium (concentrations as described above) containing 30 mM glucose. The flow rate was adjusted to 0.5 ml min⁻¹ and 1-min fractions were collected directly into scintillation counting vials. Superfusion was continued for 10 min with plain medium to establish a

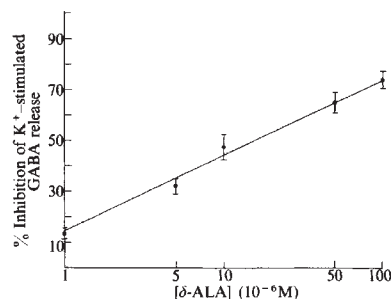


Fig. 1 A dose-response curve for the inhibition of potassium-induced GABA release by δ -ALA. The horizontal axis gives concentration of δ -ALA ($\times 10^{-6}$ M) on a logarithmic scale. Represented on the vertical axis is the % inhibition of potassium-stimulated (55 mM) GABA release. Each point is the mean % inhibition $\pm$ s.e.m. ($n = 4$). The curve is the best fit to the data by the method of least squares ($r^2 = 0.97$). It was not possible to produce 100% inhibition of the stimulated release as δ -ALA alone at concentration $> 10^{-4}$ M caused increased efflux of radioactivity from the synaptosomes³.

Table 1 Drug effects on the potassium-stimulated release of GABA from preloaded synaptosomes

Drug present	Induced release (% increase over baseline control)	Mean drug effect on induced release
55 mM K ⁺ alone (6)	188 ± 26.9*	
10 ⁻⁴ M δ -ALA alone (4)	23 ± 9.9	
55 mM K ⁺ + 10 ⁻⁴ M δ -ALA (4)	48 ± 6.4†	-74.5%
55 mM K ⁺ + 10 ⁻⁴ M δ -ALA + 10 ⁻⁶ M bicuculline (4)	191 ± 13.2	+1.6%
55 mM K ⁺ + 10 ⁻⁵ M δ -ALA (4)	92 ± 13.5†	-51.1%
55 mM K ⁺ + 10 ⁻⁵ M δ -ALA + 10 ⁻⁶ M picrotoxin (4)	187 ± 10.2	-0.5%
55 mM K ⁺ + 10 ⁻⁶ M δ -ALA (4)	163 ± 7.5‡†	-13.0%
55 mM K ⁺ + 10 ⁻⁶ M δ -ALA + 10 ⁻⁶ M bicuculline (4)	197 ± 13.1	+4.8%

Experiments were carried out as described in the text. The values in parentheses indicate the number of times the experimental conditions in the first column were repeated. Baseline efflux was calculated as % total tissue stores released per fraction ($0.57 \pm 0.068\%$, $n = 7$). The second column refers to the % increase in efflux over this control baseline level; in each case the mean $\pm$ s.e.m. is given. Statistical significance was calculated using the two-sided Student *t*-test. The third column represents the average drug-induced change compared with the release caused by 55 mM potassium alone. δ -ALA at concentrations $<10^{-4}$ M had no effect on the basal release of radioactivity from the synaptosomes. Bicuculline and picrotoxin alone (10^{-6} M or 10^{-5} M) did not facilitate the release of GABA, presumably because the flow of superfusion fluid removed any GABA as it was released. In the absence of negative feedback effects to antagonise, the GABA receptor antagonists has no effect on basal efflux.

* Substitution of Mg²⁺ for Ca²⁺ in the medium did not significantly affect basal release of radioactivity, although $76 \pm 2.2\%$ ($n = 4$) of the K⁺ (55 mM) induced release was Ca²⁺ dependent.

† $P < 0.001$, significant difference from 55 mM K⁺ alone.

‡ $P < 0.05$, significant difference from 55 mM K⁺ alone.

baseline rate of GABA efflux; the plain medium was then exchanged for the test medium containing the drugs listed in Table 1 and superfusion continued for a further 15 min. Aquagel I liquid scintillant, 5 ml, was added to each vial and the radioactivity of the vials was counted in a Packard Tri-Carb spectrometer at 37% efficiency. Filters were solubilised in 10 ml of Aquagel I and counted similarly. Stimulation of GABA release was calculated as the percentage increase in efflux over the baseline unstimulated level. The data obtained were analysed using a two-sided Student *t*-test and the results are shown in Table 1.

δ -ALA dose dependently reduced the potassium depolarisation-induced release of GABA from the nerve endings (Fig. 1). The reduction was significant at a δ -ALA concentration of 10^{-6} M, and at a concentration of 10^{-4} M 75% of the stimulated release was abolished. The concentration of δ -ALA which inhibited 50% of the potassium-stimulated release was approximately 10^{-5} M. In all cases the reduction was prevented by the GABA receptor antagonists bicuculline or picrotoxin (10^{-6} M) (Table 1). δ -ALA at 10^{-6} M or 10^{-5} M exhibited no effects on the basal release of GABA although, at 10^{-4} M, GABA efflux was slightly stimulated (Table 1). Neither bicuculline nor picrotoxin alone (10^{-6} M or 10^{-5} M) modified the basal release of radioactivity in the conditions used here. This is expected, as we consider that any GABA released from the tissue is rapidly removed from the superfusion chamber by the flow of medium. The concentration of GABA in the medium does not therefore reach the levels necessary to stimulate presynaptic receptors and hence the bicuculline or picrotoxin has no negative feedback effects to antagonise.

The evidence presented shows that δ -ALA, an ω -amino acid with a five-carbon chain similar in structure to GABA, reduces the amount of radioactivity released as a result of potassium stimulation, and that this effect is prevented by the GABA

receptor antagonists bicuculline and picrotoxin. This is consistent with the idea that δ -ALA acts as an agonist at presynaptic receptors in the mammalian central nervous system. Indeed, it seems to be a fairly powerful agonist, inhibiting 50% of the stimulated release at a concentration of 10^{-5} M.

Although the influx of δ -ALA into the brain is limited⁸, it has been demonstrated that δ -ALA can penetrate the blood-brain barrier of the rat at plasma concentrations known to occur in acute intermittent porphyria and appear in the cerebrospinal fluid in micromolar concentrations¹⁵. In addition, cerebrospinal fluid levels of 1.6 – 2.0×10^{-5} M have been recorded in patients during acute attacks characterised by hyperexcitability and convulsions¹². Our findings strongly support a direct mechanism whereby δ -ALA acts to inhibit GABA release from inhibitory nerve endings. This effect is evident *in vitro* at concentrations as low as and lower than those already recorded in the central nervous system during acute porphyric episodes.

This work was supported in part by grants from the South African MRC and the Atomic Energy Board of South Africa. The authors are Mellor Research Fellows.

M. J. W. BRENNAN
R. C. CANTRILL

Department of Medical Biochemistry,
The Medical School,
University of the Witwatersrand,
Johannesburg 2001,
South Africa

Received 14 May; accepted 20 June 1979.

1. Eales, L. S. *Afr. J. Lab. clin. Med.* **9**, 151–162 (1963).
2. Tschudy, D. P. in *Duncan's Diseases of Metabolism* (eds Bondy, P. K. & Rosenberg, L. E.) 775–824 (Saunders, Philadelphia, 1974).
3. Brennan, M. J. W. & Cantrill, R. C. *J. Neurochem.* **32**, 1781–1786 (1979).
4. Brennan, M. J. W. & Cantrill, R. C. *J. Neurochem.* (in the press).
5. Kramer, S., Becker, D. & Viljoen, D. S. *Afr. med. J.* **47**, 1735–1738 (1973).
6. Meyer, U. A. & Schmid, R. in *Brain Dysfunction in Metabolic Disorders* (ed. Plum, F.); *Res. Publ. Ass. Res. nerv. ment. Dis.* **53**, 211–223 (1974).
7. Becker, D. M., Viljoen, J. D. & Kramer, S. *Biochim. biophys. Acta* **225**, 26–34 (1971).
8. Shanley, B. C., Neethling, A. C. Percy, V. A. & Carstens, M. S. *Afr. med. J.* **49**, 576–580 (1975).
9. Snodgrass, S. R. *Nature* **274**, 392–394 (1978).
10. Mitchell, P. R. & Martin, I. L. *Nature* **274**, 904–905 (1978).
11. Brennan, M. J. W., Cantrill, R. C. & Epstein, H. *4th natn. Congr. S. Afr. biochem. Soc.*, 102 (1979).
12. Sweeney, V. P., Pathak, A. M. & Asbury, A. K. *Brain* **93**, 369–380 (1970).
13. Brennan, M. J. W. & Cantrill, R. C. *J. Neurochem.* **31**, 1339–1341 (1978).
14. Cantrill, R. C. & Brennan, M. J. W. *Experientia* (in the press).
15. McGillion, F. B., Thompson, G. G., Moore, M. R. & Goldberg, A. *Biochem. Pharmacol.* **23**, 472–474 (1974).

Correlation of local stability of DNA during melting with environmental conditions

DNA MELTING studies have revealed fine structural features in the melting profiles of many natural DNAs^{1–4}. Restriction fragments have been used to assign modes showing up in melting profiles to the denaturation of well delineated segments inside the DNA molecule^{1,3} and possible correlations of these segments with biological functions have been suggested^{1,5,6}. However, when the exact sequence of natural DNAs became available, known genetic or regulatory features extending over tens to hundreds of base pairs did not appear to be characterized systematically by simple physical parameters. For example, the base pair composition fluctuation of Φ X174 DNA is almost smooth^{6,7}, whereas several distinct modes are seen in experimental melting profiles^{8,9}, each reflecting separate melting of segments of hundreds of base pairs. Attempts have been made to calculate the melting profiles of DNA from its sequence^{10–13}, and two calculated melting profiles have been published for Φ X174 DNA. Both calculated curves show fine structure, but comparison with the experimental profiles showed 'no obvious

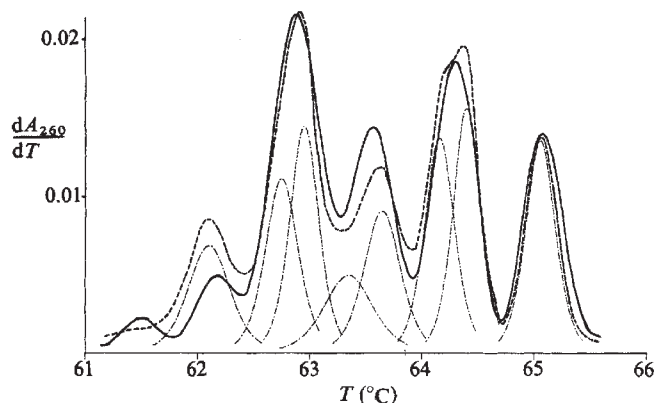
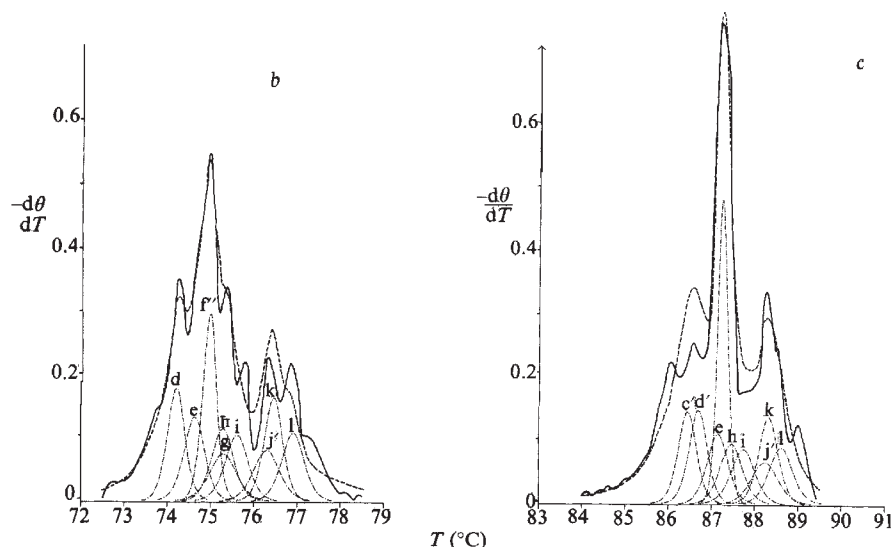
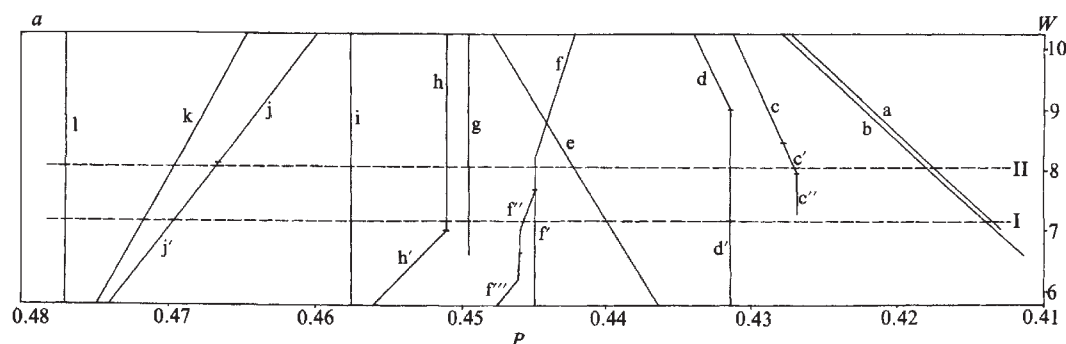


Fig. 1 Melting profile of a restriction fragment of SV40 DNA, extending from base pair 1,967 to 3,409 (origin: *EcoRI* restriction site). Solid line: experimental melting profile, obtained in 10 mM Na⁺ (citrate), 10⁻⁴ EDTA; temperature incremented stepwise by 0.05 °C, mean heating rate 4 °C per h (see ref. 1 for experimental setup and procedure, and ref. 19 for data processing). Dashed line: calculated melting curve, for the sequence published by Fiers *et al.*²⁰. Parameters used: $T_{AT} = 44.8$ °C, $T_{GC} = 93.9$ °C, $W = 4.07$, $b = 1.114$ (loop entropy factor), $\beta = 0.034$ (mode form factor); see text and refs 14 and 16 for definition of these parameters. Dashed-dotted curves correspond to the melting profiles of individual, cooperatively melting segments (only segments with 100 or more base pairs are shown).

Fig. 2 a, Part of the phase diagram of Φ X174 DNA computed from its sequence²¹. Ordinate: $W = V/(f_1 - f_0)$, abscissa: $p = (T - T_{AT})/(T_{GC} - T_{AT})$ (see text and ref. 16). The lines divide domains having different composition of melted and helical segments in the ground state. Neighbouring domains differ by the state of a given stretch of base pairs (only domains differing by the state of at least 200 base pairs are shown). The base pair sequences changing state on crossing the boundaries are: (letters refer to letters drawn in the diagram; numbers indicate base pairs, the origin being the end of gene H (ref. 21)).



b, Solid line: experimental Φ X174 DNA melting profile, 30 mM Na⁺ (from ref. 8). Dashed line: computed curve corresponding to intercept I ($W = 7.17$) of the phase diagram. Other parameters: $T_{AT} = 55.4$ °C, $T_{GC} = 104.5$ °C, $\beta = 0.084$, $b = 1.114$. This curve differs slightly from that obtained by Azbel¹⁴ because we have used a somewhat different, more accurate sequence²¹. Dashed-dotted curves: melting profiles of individual segments (only the most prominent modes are depicted). **c**, Full line: experimental Φ X174 melting profile, 195 mM Na⁺ (from ref. 8). Dashed line: computed curve, corresponding to intercept II ($W = 8.02$) of the phase diagram. Other parameters: $T_{AT} = 68.1$ °C, $T_{GC} = 111.2$ °C, $\beta = 0.081$, $b = 1.114$. Dashed-dotted curves as in **b**.

systematic relationship' (ref. 8). The very different profiles observed for the same Φ X174 DNA in two different ionic strengths⁸ could not be accounted for by the theory used. Thus the melting profile is not determined by the DNA sequence alone. However, recently good agreement between the calculated and experimental melting profile of Φ X174 DNA was obtained by Azbel¹⁴⁻¹⁶. Azbel's model is based on the observation that nucleation of the helix-coil transition requires much higher energies than that exchanged between base pairs, so that behaviour of DNA is mainly correlated to its ground state at a given temperature inside the melting range (see ref. 16). We have applied Azbel's model to another DNA of known sequence (a restriction fragment of SV40 DNA) and Φ X174 DNA in two different environmental conditions, and we obtained good agreement between predictions and experimental observations.

It can be seen from Figs 1 and 2b, c that the agreement between calculated melting profiles and experimental curves, obtained in this laboratory or published by Vizard *et al.*⁸, is satisfactory both in terms of the number of modes and the shapes of the various profiles. Minor differences (most are within the experimental error bar) may be due to linear, first order approximation to Azbel's model.

An interesting aspect of the model is the fact that a phase diagram can be constructed to show the variation of the ground state as a function of two dimensionless parameters which are related to standard thermodynamic functions. The phase diagram shown in Fig. 2a predicts the state of Φ X174 DNA, inside its melting range, for a range of $W = V/(f_1 - f_0)$, where V is the boundary energy between helix and coil regions, and $f_1 - f_0$

is the difference in free energy between AT and GC pairs, the abscissa of the phase diagram is given by $p = (T - T_{AT}) / (T_{GC} - T_{AT})$ (ref. 16).

As W varies, the size and behaviour of some of the segments remain invariant, such as segment 405–744. Others change, however, such as segment 2,880–3,678 which changes its state at $p = 0.445$ when $W \geq 7.71$, is split into two segments when $W < 7.71$; segment 3,451–3,678 changes state at $p = 0.445$ as above, and 2,880–3,451 changing state at higher p (temperature). Thus size, location, and relative stability of homostable (cooperatively melting) segments inside DNA may be modified by changes in W . These can be induced by various environmental changes. A complete catalogue of the solution parameters acting on W is not yet established, but experimental evidence shows that ionic strength^{17,18}, pH and type of counterion¹⁸ act on the thermodynamic functions entering W and can modify the melting profile of a given DNA.

Extending the phase diagram concept to DNA *in situ*, we may consider changes in W induced by the interaction of DNA with metabolites, activators, histones, polyamines, and so on. These could modify the size, location and relative stability of homostable cooperatively-behaving segments inside the DNA. Such stability switching and shifting could account, at least in part, for the regulatory activity of these agents. Attempts to correlate the complete phase diagram with known genetic features are under way.

We thank Professor W. Fiers for samples of SV40 DNA and Professors P. G. de Gennes and M. Ya'Azel (who sent us his manuscript before publication) for useful discussions. J.G.A. was in receipt of EMBO fellowship ALTF 118–78. Financial support from DGRST (MMR, grant 77–70296) to C.R. is acknowledged.

J. GABBARO-ARPA*
P. TOUGARD
C. REISS*

CNRS-Form. 031835,
91 Gif sur Yvette, France

* Present address: IRBM, Tour 43, Université Paris VII, 2 Place Jussieu, 75005 Paris, France.

Received 12 February; accepted 25 June 1979.

- Reiss, C. & Gabbarro-Arpa, J. *Prog. mol. subcell. Biol.* **5**, 1–30 (1977).
- Vizard, D. L. & Ansevin, A. T. *Biochemistry* **15**, 741–750 (1976).
- Michel, F. J. *molec. Biol.* **89**, 305–326 (1974).
- Lyubchenko, Y. L., Frank-Kamenetskii, M. D., Vologodskii, A. V., Lazurkin, Y. & Gautze, G. C. Jr. *Biopolymers* **15**, 1019–1036 (1976).
- Tachibana, H., Wada, A., Gotoh, O. & Takamami, M. *Biochim. biophys. Acta* **517**, 319–328 (1978).
- Tong, B. Y. & Batterby, S. J. *Biopolymers* **17**, 2933–2937 (1978).
- Ueno, S., Tachibana, H., Husimi, Y. & Wada, A. *J. Biochem.* **84**, 917–924, (1978).
- Vizard, D. L., White, R. A. & Ansevin, A. T. *Nature* **275**, 250–251 (1978).
- Lazurkin, Y. S. *Molec. Biol.* **11**, 1007–1017 (1978).
- Tong, B. Y. & Tong, S. Y. *J. chem. Phys.* **54**, 1317–1324 (1971).
- Poland, D. *Biopolymers* **13**, 1859–1871 (1974).
- Fixman, M. & Freire, J. J. *Biopolymers* **16**, 2693–2704 (1977).
- Frank-Kamenetskii, M. D. & Frank-Kamenetskii, A. D. *Molec. Biol.* **3**, 375–383, (1969).
- Azbel, M. Y. *Proc. natn. Acad. Sci. U.S.A.* **76**, 105, (1979).
- Azbel, M. Y. *Phys. Rev. Lett.* **31**, 589–593 (1973).
- Azbel, M. Y. *Ann. Isr. Phys.* **2**, 26–36 (1978).
- Gruenwedel, D. W. *Biochim. biophys. Acta* **340**, 16–30 (1974).
- de Zamaroczy, M. thesis, Univ. Paris (1978).
- Gabbarro-Arpa, J. *Analyt. Biochem.* **91**, 308–322 (1978).
- Fiers, W. *Nature* **273**, 113–120 (1978).
- Sanger, F. et al. *J. molec. Biol.* **125**, 225–246 (1978).

Environmental factors affecting the quantity of alcohol dehydrogenase in *Drosophila melanogaster*

THE alcohol dehydrogenase gene in *Drosophila melanogaster* has recently been the subject of intense study because it has properties that make it almost ideal for the investigation of genetic fine structure and control in a higher eukaryote^{1–3}, and for research on the population genetics of enzyme polymorphism^{4,5}. There have been reports of genetic variation in

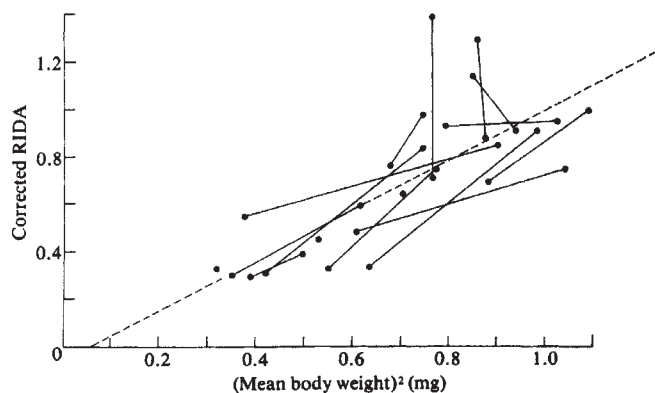


Fig. 1 The regression of corrected radial immunodiffusion area (RIDA) on the square of the mean body weight in batches of 25 4-day old male flies. The dotted line represents the overall regression ($y = 1.045x - 0.051$, $P < 0.001$). Measurements made on the same inbred line in different generations are joined by continuous lines. Inbred lines on which only single measurements were made appear as disconnected spots. The corrected RIDA is the area of the sample divided by the area of a standard extract (a mass homogenate of flies from a white-honey stock, used throughout the experiment) measured on the same gel. The flies came from 15 inbred lines, all homozygous for the *Adh*^s allele, derived from the Edinburgh Kaduna population (established with flies collected in Nigeria in 1949 and maintained since then in population cages). The lines were grown in one-third-of-a-pint milk bottles containing dead-yeast-cornmeal-molasses medium (11 g of yeast per l of vkm) at 25 °C. Adults were isolated as they emerged from pupae. 25 4-day old males were homogenised in phosphate buffer (pH 8.0), and the homogenates were passed through a Millipore filter. 4- μ l samples of filtrate were introduced into the wells of a 1% agarose gel (pH 8.6) containing 2% rabbit antiserum against homogenised *D. melanogaster*. The gels were incubated for 48 h at 25 °C and the precipitin rings stained for ADH with nitroblue tetrazolium-methylene blue¹³. Each point represents the mean of two replicate gels.

alcohol dehydrogenase activity within and between electrophores^{5–10}, and of a putative control mutant with half the 'normal' activity¹¹. The studies described here show that environmental factors can change the amount of ADH protein in a fly by at least fourfold. They invalidate conventional methods of correcting for body size in studies of the activities and amounts of enzymes in *Drosophila*, and therefore cast doubt on much recent work. They also raise the possibility of an inducible eukaryotic enzyme that is suitable for refined genetic analysis.

The first observation, originally seen in experiments carried out by P. C. Hillier, is that both the activity of the enzyme (measured spectrophotometrically) and its quantity (assayed by single radial immunodiffusion¹²) can rise disproportionately with body weight. This is illustrated in Fig. 1, which shows the relationship between the area of the precipitin ring (strictly proportional, in our gels, to the quantity of enzyme) and body weight in a series of inbred lines homozygous for the *Adh*^s (slow) allele derived from the Kaduna population¹³. There is a fourfold range in the quantity of enzyme in flies from different cultures. That this is not due solely to genetic variation between the inbred lines is shown by the large differences between generations in one line. Indeed, there is no evidence that the various SS lines derived from Kaduna differ genetically in their quantities of ADH. The differences are not due to errors of measurement; the control areas (from the standard extract) are relatively constant from gel to gel, and the average difference between corrected replicate measurements is 11.5%.

The increase in the quantity of ADH with body weight is not linear. In the Kaduna lines it rises proportionally to the square of the body weight (fitting the power curve $y = ax^b$ to the data gives $b = 2.24$). Figure 2 shows that this phenomenon can occur within one inbred line. It gives measurements of single male flies from the Montgomery strain. In this case the amount of ADH is

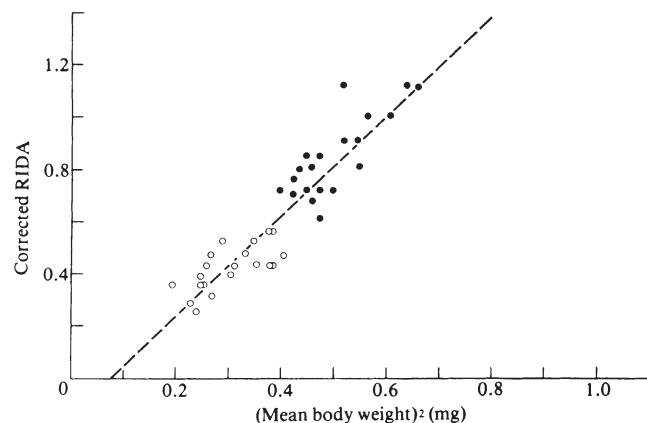


Fig. 2 The regression of corrected RIDA on the square of body weight in single 4-day old male flies from the Montgomery inbred line (homozygous for the *Adh^S* allele). The dotted line represents the overall regression ($y = 1.915 - 0.144x$, $P < 0.001$) which is estimated for comparison with Fig. 1. The measurements were made in two separate generations, represented by open and closed circles. The difference between the generations, due to uncontrolled environmental differences, is very clear. The methods were identical to those given in Fig. 1 legend, except that individual flies were homogenised, and the homogenates centrifuged cold at 13,000g for 2 min, and that the gels contained 3% sheep antiserum. Each point represents the mean of two replicate gels.

approximately proportional to the cube of the weight ($b = 3.13$), and once more there is a fourfold range. The significant difference between the slopes of the regressions shown in Figs 1 and 2 suggests that the lines differ in their responses to the environmental factor that increases the quantity of ADH, but this remains to be confirmed.

Because the weight of adult flies is known to be influenced by the conditions of larval growth, we introduced newly hatched first instar Kaduna *SS* larvae into 2.5-cm shell vials containing 10 ml of dead-yeast-cornmeal-molasses medium with various quantities of dead yeast. Four vials had 50 g of yeast per l of medium, four had 10 g, four had 2 g and four had no yeast at all. The quantities of ADH in single 4-day-old adult males that emerged from these vials were measured by radial immunodiffusion. The results, given in Table 1, show that although the progressive addition of yeast increased the mean body weight by about 30%, it more than doubled the average quantity of ADH. Within the treatment groups there was only a slight association between body weight and ADH. It did not depart from linearity, and was significant in only one group (10 g yeast per l). Between the groups, however, the association is very striking, and the differences are highly significant ($P < 0.001$).

Table 1 The effects of different concentrations of dead yeast on the amount of ADH protein

Amount of dead yeast (g per l medium)	No. of vials	No. of 4-day male flies measured	Mean weight mg (s.e.)	Mean RIDA mm ² (s.e.)
0	4	18	0.719(0.015)	30.13(1.37)*
2	4	17	0.791(0.021)	47.88(2.89)*
10	4	18	0.901(0.015)	81.83(2.02)
50	4	20	0.906(0.018)	80.70(1.73)

1st instar larvae were obtained from eggs deposited by Kaduna *Adh^S* homozygotes on laying dishes, and were grown in 2.5-cm shell vials containing 10 ml of medium (50 larvae per vial) at 25 °C. The methods were otherwise identical to those given in Fig. 2 legend. The asterisks represent means based on RIDAs that were significantly heterogeneous between vials ($P < 0.05$). Despite this heterogeneity the differences in RIDA between treatments are highly significant ($P < 0.001$).

These findings have several implications. The amount of enzyme is 'allometric' with body weight, rather than strictly proportional to it. This means that conventional ways of correcting for body size (dividing the activity or quantity of enzyme either by the body weight itself or by the total soluble protein, which is proportional to body weight) may give seriously misleading results. Supposed genetic differences in the activity or quantity of enzyme may merely reflect persistent disparities of body weight and/or the conditions of culture. This uncertainty applies to the report of a twofold difference in the quantity of ADH between genotypes⁷, and to the recent claim that a 30% difference between a line selected for resistance to ethanol and an unselected control was due to the action of 'regulatory' genes¹⁴. It may also apply to work on other *Drosophila* enzymes, for which the effects of variations in the medium have not been studied. Our results suggest that screening for activity mutants may be greatly improved by a careful control of the medium. We note that disparities between lines in the conditions of culture may seem to be genetic because micro-organisms can be passed on from the medium of the parents to that of their offspring.

The causes of the changes in the quantity of ADH are unknown. They could be due to an increased synthesis of enzyme in favourable conditions, to a decreased breakdown, or to changes in the time course of synthesis or breakdown. These possibilities are being studied, but in any event there is evidence that the precipitating factor is some part of the larval environment correlated with the quantity of yeast. A constituent or product of the yeast is the most obvious candidate. Because the flies were raised on dead yeast the factor is unlikely to be ethanol. Two per cent ethanol added to medium containing 2 g yeast per l does not increase the quantity of ADH (unpublished observation). If the changes prove to be a case of induction, and if the inducing substance can be identified, these observations will offer interesting possibilities for the study of eukaryotic gene control.

We thank the MRC and SRC for financial support.

BRYAN CLARKE
ROBERT G. CAMFIELD
ALISON M. GALVIN
CHRISTOPHER R. PITTS

Genetics Research Unit,
Queen's Medical Centre,
Clifton Boulevard, Nottingham, UK

Received 4 May; accepted 5 July 1979.

1. Sofer, W. H. & Hatkoff, M. A. *Genetics* **72**, 545 (1972).
2. Schwartz, M. & Sofer, W. *Genetics* **83**, 137 (1976).
3. Schwartz, M. & Jörnval, H. *Eur. J. Biochem.* **68**, 159 (1976).
4. Clarke, B. *Genetics* **79**, 101 (1975).
5. Morgan, P. *Heredity* **34**, 124 (1975).
6. Ursprung, H., Smith, K. D., Sofer, W. H. & Sullivan, D. T. *Science* **160**, 1075 (1968).
7. Gibson, J. B. *Experientia* **28**, 975 (1972).
8. Birley, A. J. & Barnes, B. W. *Heredity* **31**, 413 (1973).
9. Day, T. H., Hillier, P. C. & Clarke, B. *Biochem. Genet.* **11**, 155 (1974).
10. Ward, R. D. *Biochem. Genet.* **12**, 449 (1974).
11. Thompson, J. N., Ashburner, M. & Woodruff, R. C. *Nature* **270**, 363 (1977).
12. Mancini, G., Carbonara, A. O. & Heremans, J. F. *Immunochimistry* **2**, 235 (1965).
13. Johnson, F. M. & Denniston, C. *Nature* **204**, 906 (1964).
14. McDonald, J. F., Chambers, G. K., David, J. & Ayala, F. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4562 (1977).

Erratum

In the letter 'Influence of leaves on sporophore production by fungi forming sheathing mycorrhizas with *Betula* spp.' by F. T. Last *et al.*, *Nature* **280**, 168–169, in paragraph 2 line 5, for '*Betula pendula* collected by Roth', read '*Betula pendula* Roth collected from ...'.

Corrigendum

In the letter 'Non-selective isolation of human somatic cell hybrids by unit-gravity sedimentation' by P. L. Chang *et al.*, *Nature* **278**, 168–169, in paragraph 4 lines 9–10 the letters A and B referring to Fig. 1 have been transposed. It should read '... 2.4 units (A) ... compared to 4.1 units (B) ...'.

reviews

Selection in the social insects

Brian Charlesworth

Caste and Ecology in the Social Insects. By G. F. Oster and E. O. Wilson. Pp. 352. (Princeton University Press: Princeton, New Jersey, and Guildford, UK, 1979.) Hardback £12.50; paperback £4.70.

In the *Origin of Species*, Darwin wrote: "... I believe that natural selection, by acting on the fertile ants ... could form a species which should regularly produce neuters, all of large size with one form of jaw, or all of small size with widely different jaws; or lastly, and this is the greatest difficulty, one set of workers of one size and structure, and simultaneously another set of workers of a different size and structure". He attributed this to "family selection": "slight modifications of structure or of instinct, correlated with the sterile condition of certain members of the community, have proved advantageous: consequently the fertile males and females have flourished, and transmitted to their fertile offspring a tendency to produce sterile members with the same modifications". Apart from the brief formulations of kin selection by Fisher and Haldane in the 1930s, relatively little was added to this wonderful insight of Darwin's about the evolution of the social insects until W. D. Hamilton's seminal papers in the early 1960s. Since then, interest in this fascinating subject has steadily grown.

Oster and Wilson's book is a synthesis of recent work, with particular emphasis on the problem of the selective factors affecting the numbers and kinds of sterile castes. Professor Oster is well-known for his outstanding contributions to mathematical ecology, and Professor Wilson for his leadership in the field of sociobiology. One would therefore expect this book to be a formidable combination of theoretical skills and wide knowledge of the natural history of the social insects. This expectation is certainly fulfilled. The book alternates between theoretical analyses and discussions of the relevant biological facts. Although the mathematical derivations are relegated to appendices, I suspect that many biologists will, like myself, find the theory rather heavy going. Nonetheless, many important questions are discussed, and workers on the evolution of social behaviour will need to study this book with care.

Chapter 1 provides an introduction to the problem of explaining the varied behavioural repertoires of the workers,

and (in some groups) the existence of distinct physical castes. The authors state their intention of attacking the problem mainly through the concept that "castes and division of labour are subject to optimisation by natural selection at the colony level". (By colony-level selection, they mean Darwin's family selection; curiously enough, Darwin's contribution is never mentioned.)

Chapter 2 reviews the extensive data on different types of colony life-cycles. Some interesting suggestions are put forward concerning the selective factors which may influence the ways in which colonies can be founded (swarming as opposed to foundation by reproductives only, single as opposed to multiple queens, and so on), and there is a detailed theoretical analysis of the optimal allocation of resources between workers and reproductives during the development of a colony. This seems to provide a convincing explanation for the fact that, in temperate-zone species, reproductives are usually produced as a burst towards the end of the season, with the production of workers being simultaneously terminated. To my mind, this is the most successful application of the theory to a real-life problem in the whole book, as the theoretical predictions are not intuitively obvious, but follow from a plausible model. A note of caution is in order, however, as recent theoretical work suggests that the frequency of workers within a colony may be subject to control by kin selection or parental manipulation. A colony-level selection approach to this problem may not, therefore, be entirely satisfactory.

Chapter 3 is a discussion of the genetic basis of selection on social insects, particularly in relation to the sex-ratio problems originally treated by Trivers and Hare. I found this chapter somewhat disappointing. It assumes that selection should act to maximise either the fitness of the colony, the inclusive fitness of workers, or the inclusive fitness of the queen. It is not at all clear what the relevance of the colony's fitness is to a question such as the optimal allocation of resources between male and female reproductives, as evolutionary change in this variable must proceed by the substitution of genes which express themselves in individuals (queens or workers). Furthermore, considerable algebraic contortions have to be gone through in

order to define the fitness parameters correctly, and to obtain solutions to the maximisation problems. A simpler approach would have been to use the standard population genetics methods for determining the values of the population strategies which are immune to invasion by rare genes causing deviant strategies — Maynard Smith's concept of the Evolutionarily Stable Strategy (ESS). These methods have been successfully applied to the sex-ratio problem by E. L. Charnov and others, and have the advantages both of simplicity and of being based firmly on gene frequency analysis. Chapter 3 also contains a rather misleading account of the ESS concept, which conveys the impression that an ESS is equivalent to a Nash equilibrium in standard games theory. But in the animal conflict models originally studied by Maynard Smith, members of a population playing a strategy S_i have the fitness associated with contests between pairs of S_i individuals. A rare mutant playing a different strategy S_j will, to a good approximation, have the fitness associated with contests between S_j and S_i . S_i is an ESS if the former fitness exceeds the latter for all values of S_j . This has nothing in common with a Nash equilibrium, which involves a pair of opponents with different interests. Nash equilibria may be relevant to situations where different members of the population have conflicting interests (for example, the queen-worker conflict studied by Oster and Wilson), but the ESS concept is broader.

Chapter 4 gives a lucid account of the major facts about temporal and physical castes. Chapter 5 outlines the concepts underpinning the theoretical models of caste evolution developed in later chapters, the central idea being that of the "caste distribution function" or CDF, which specifies the distribution of the numbers of workers with respect to physical and behavioural caste, as a function of time, throughout the colony's development. The CDF is assumed to be adjusted by colony-level selection such that the total number of reproductives is maximised. Chapter 6 develops this model, using dynamic optimisation theory to solve for the optimal CDF. As the authors make clear, the problem is not so much why there should be distinct castes, but why there are usually so few. They claim that, in

general, a size distribution of workers which precisely matches the distribution of size of available resource items is not to be expected. Their argument relies heavily on the concept of "fidelity cost", which assumes that an abrupt increase in the number of individuals in a given class creates an energy cost to the colony. This is an interesting idea, but seems rather difficult to evaluate empirically, in view of the necessarily arbitrary way in which the fidelity cost is introduced into the models. Chapter 7 is concerned with foraging strategies of workers. Chapter 8 is a useful discussion of the role of

optimisation theory in evolutionary biology, and Chapter 9 is a brief survey of unsolved problems. As the authors make clear, both here and in the preface, the major factors determining the evolution of the detailed features of caste ratios and behavioural flexibility remain largely unknown. Nevertheless, their use of quantitative models highlights the problems very clearly and should prove very stimulating to research in this area. □

Brian Charlesworth is Lecturer in Biology at the University of Sussex, Brighton, UK.

Affinity chromatography

An Introduction to Affinity Chromatography. By C.R. Lowe. Pp. 253. (North-Holland: Amsterdam, New York and Oxford, 1979.) \$29.25; Dfl60.

THIS volume is part of a series on laboratory techniques. The first four chapters consist of an introduction to chromatography, and accounts of basic concepts, chemistry and techniques of affinity separations. Chapter 5 is a selection of applications (protein-enzyme area), and a separate chapter covers regulatory macromolecules and more complex structures (although how antibody-antigen systems fit this category escapes me). The last two chapters are concerned with analytical applications and special techniques (very useful). The book covers most of the earlier work in the field: 9 references out of the 240 are 1977 or later.

The author has many thoughtful remarks to make for users of affinity chromatography. Rather than present a flawless up-to-the-minute review, Dr Lowe provides a valuable set of guidelines which should inspire rather than weary the reader. Because the technique is developing so rapidly some parts of the book are incomplete. Thus, it is now clear that triazine dyes are widely used as affinity ligands, their main advantage being the combination of capacity and selectivity. A slightly misleading comment on p455 suggests that all albumins will bind to Cibacron Blue columns when this property refers to human albumin only.

In the next edition, I would like to see some of the following recent advances discussed: the receptor purification advocated by Baulieu, involving alteration of the molecular weight of a protein using soluble dextran-ligand conjugates, electrophoretic desorption of high affinity biospecific adsorbents, light-reversible affinity ligands, and high performance affinity chromatography.

The sections on ligand concentration, temperature effects, the importance of

ligand preassembly and the concept of group specific affinity matrices did not do justice to the author's work in these areas. I would also have liked to see a warning of the toxicity of S-trichlorotriazine (cyanuric chloride).

Pride of place in accounts of affinity chromatography goes to the now famous use of immobilised oxamate to adsorb lactate dehydrogenase in the presence of NAD(H). Dr Lowe's book leaves this bastion of theoretical preconception unassaulted, although Wood *et al.* have convincingly shown (1977) that replacing CNBr with triazine chemistry abolishes the biospecificity of the oxamate column.

Apart from the above minor criticisms, I found the balance between commercials, laboratory methods (for which Dr Lowe has an obvious flair) and theory enjoyable and well thought out.

P.D.G. Dean

P.D.G. Dean is Senior Lecturer in the Department of Biochemistry at the University of Liverpool, UK.

Biological semiconduction

Dielectric and Electronic Properties of Biological Materials. By R. Pethig. Pp.376. (Wiley: Chichester, UK, and New York, 1979.) £15.

THE possibility that electronic energy bands exist in certain solid-state biological systems, playing a role in energy migration (Möglich and Schön; Jordan 1938) and electron transfer (Szent-Györgyi, 1941) has been around for over 40 years. During this time there has occurred a small but steady amount of theoretical and experimental progress on biological semiconduction, that is, on the electron transfer problem. Earlier work tended to show the conductivities were too low for biological significance, but with passage of time, the effects of hydration, and of charge-transfer, in inducing appreciable electronic conduction in proteins, all go towards increasing the prospects of a role of semiconduction in certain biological processes. The recent demonstration of high conductivity in reduced cytochrome c_3 by

Inokuchi and coworkers, is another step in the same direction. Dr Pethig's book is aimed at introducing newcomers from a variety of disciplines to the field, and these are to be welcomed, as the need to increase the ratio of facts to speculations is rather urgent.

Very sensibly Pethig's first three chapters deal successively with dielectric theory, the intrinsic dielectric properties of biopolymers, and the effects of solvation on the dielectric properties. In these he established the basic solution properties, and then goes on in his fourth chapter to the effects of water on the solid-state behaviour, dealing with X-ray studies and water adsorption. In his fifth chapter Pethig outlines interfacial dielectric phenomena, the Maxwell-Wagner effect and its developments. Although this treatment is clear, I believe it could usefully have been a greater length, as the separation of Maxwell-Wagner from dipolar and hopping loss mechanisms is a fundamental problem in this field. Per contra, the sixth chapter dealing with dielectrophoretic studies is mainly of interest in connection with preparative separation methods, and is slightly off the main theme of the book. Chapter 7 deals with biological membranes and tissues, with brief notes on black lipid membranes, and includes three pages on the sometimes controversial effects of electromagnetic fields. With the final two chapters, 8 and 9, we come to the main part of the book, in which the author's own research interests are to be found. Chapter 8, on electrons, energy levels and energy bands, provides the basic background for crystalline and amorphous semiconductors, and includes an account of band gap and band width calculations for biopolymers, from the pioneer work of M.G. Evans and J. Gergely (1949) to the present. Chapter 9 sets out the experimental data on semiconductivity, Hall effects, charge-transfer interactions, piezoelectric effects and so on for proteins, DNA and other biological molecules. Photoconductivity is mentioned, but rather too briefly in view of its possible importance in vision and photosynthesis. Chapters 8 and 9, however, could scarcely be comprehensive in the space available, and Pethig's approach has been to take a few selected topics in depth. So, for example, there is a thoughtful appraisal of the effects of adsorbed water on conductivity, and an account of Pethig's own work on the methylglyoxal-protein system, in connection with Szent-Györgyi's recent suggestions concerning carcinogenesis.

I believe this book gives an attractive introduction to its field. Mathematics are kept within the bounds necessary for a general readership, and in spite of the author's engineering background the emphasis is always on molecules rather than equivalent circuits. It is logically arranged, reads easily, and there is an ample supply of references. As a personal preference, I think a chapter summarising

the well-established field of ionic conduction in membranes, might have added to the value of the book, and formed a bridge to already well-established areas of electrophysiology. Where ionic and electronic conductivity exist side by side, there is a real problem in distinguishing one from the other. But Dr Pethig was probably wise not to trespass beyond his

brief of electronic properties, since this is a neglected area of biophysics at present, and I believe this book makes a valuable contribution towards remedying this neglect.

D.D. Eley

D.D. Eley is Professor of Physical Chemistry at the University of Nottingham, UK.

Photochemistry of small molecules

Photochemistry of Small Molecules. By Hideo Okabe. Pp. 431. (Wiley: Chichester, UK, 1978.) £24.55.

ADVANCES in techniques, made over the past decade or so, have greatly facilitated detection of photochemical primary products and made possible detailed studies of photodissociation dynamics. The information obtained, coupled with much new experimental data on the reactivities of ground state and excited atoms and radicals, has provided the basis for an improved understanding of photochemical processes. Dr Okabe's book has been written to cover the progress made in the gas phase photochemistry of molecules containing up to five atoms.

The contents of the book fall essentially into three parts. Two chapters deal with the principles of spectroscopy and of the primary photochemical process; the introductory section is concluded with a chapter on experimental techniques. The author has chosen here to emphasise certain topics that are of interest to himself; and they are topics of importance often ignored elsewhere. The thermal contribution to photodissociation, atomic resonance absorption and emission, and deviations from the Beer-Lambert Law are dealt with in detail, as are energy and angular distributions in photofragments, and the determination of dissociation energies. The chapter on techniques is similarly concerned with certain special aspects rather than with a comprehensive overview. The order of presentation is sometimes unexpected, with detail preceding generalisations. For example, a mathematical treatment of band intensities

is given in a section before, and unrelated to, that in which observations of intensity distributions are reported and the Franck-Condon principle stated.

The central section of the book consists of four chapters dealing with the quenching and reactions of excited atoms, and the photochemistry of diatomic to pentatomic molecules. For some 80 molecules, the spectroscopy and photochemistry are discussed in the light of information available up to July 1977. Although the huge amount of factual material presented necessarily precludes a deeply critical approach, the review is structured and readable.

A final chapter of the book is devoted to some topics related to photochemistry: isotopic enrichment, photochemical air pollution of the troposphere and stratosphere, and the photochemistry of the atmospheres of Mars, Venus and Jupiter. No more than an introduction can be given to these subjects in a total of 35 pages, but the chapter will provide a stimulus to the reader, as well as guiding him to more detailed references.

The author states that his book is aimed at the physical chemist, spectroscopist and atmospheric scientist interested in photochemistry. The level and content seem about right for this readership. A non-specialist would probably find the first chapters inadequate, and the review part too detailed, for the book to provide an introduction to photochemistry. But, for the practising photochemist, Dr Okabe has provided a most valuable reference book.

R. P. Wayne

R. P. Wayne is Lecturer in Physical Chemistry at the University of Oxford, and Dr Lee's Reader in Chemistry at Christ Church, Oxford, UK.

Asbestos sourcebook

Asbestos. Vol. 1: Properties, Applications and Hazards. Edited by L. Michaels and S. S. Chissick. Pp. 553. (Wiley: Chichester, UK, and New York, 1979.) £25.

In 1976, Dr Michaels and Dr Chissick, of the University of London, could discover no comprehensive source book on all aspects of asbestos. This first of two volumes is the result of a useful, and at

times entertaining, attempt to provide such a book. The volume covers the mineralogy, chemistry and physics of asbestos, its use, monitoring, identification, and substitute materials; and five chapters are devoted to asbestos-related diseases. There are also contributions on attitudes to asbestos, and dealing with asbestos problems.

Each chapter was commissioned with an obviously careful brief to be comprehensive, not too technical and independent of other chapters. This has led to some repetition, for example, of the different varieties of asbestos and of the

history of asbestos-related diseases. But this does not detract from any individual chapter and is only a slight annoyance to reviewers or other cover-to-cover readers.

Clearly, the brief that the editors gave themselves was to seek moderation. The contributors were carefully selected to exclude members of the asbestos industry and its declared opponents. This has worked remarkably well, as the contributions are well balanced without the extreme views often expressed on the emotive subject of the hazards of asbestos. This balance has been maintained effectively except in the chapter on attitudes to asbestos. There, the moderate discussion of attitudes of governments, industry and unions becomes more extreme itself when considering the extreme viewpoints of pressure groups and in the media. It was a pleasure to see the publication of such a positive response.

It was also a pleasure to see aired (for example, on page 385) certain of the inter-professional disputes which have been well-known among research workers, but have rarely been made public. This is a good move towards more openness in science.

In a series of commissioned chapters, not all can be of uniform standard. The introduction contains definitions (misleadingly headed Bibliography) of some terms. The first is of asbestos bodies, but the definition is of a fibre; the second contains one of those misprints prized by collectors — asbestosis is characterised by diffuse *intestinal* fibrosis — although it is argued on page 421 that the fibrosis is not "interstitial"; the radiographic classification given was issued by the International Labour Office in 1958, but the first version specifically to cover the radiographic changes related to asbestos exposure was not issued until 1971. The chapter on the clinical features of asbestos-related disease seems to treat the subject most superficially, particularly by comparison with the adjacent chapters. There were also a few other odd statements such as road death statistics are a hazard to health (page 330), and asbestos is the only known naturally occurring fibrous mineral (page 306). More seriously, there was no specific mention of the Medical Research Council as a source of information despite the many years of research on asbestos undertaken by its staff.

However, the overall impression is of a most comprehensive source book, useful to almost everybody who needs to discover information about asbestos. It goes a long way towards the editors' aims of public education. The only missing information seems to be the contents list for volume two.

Charles E. Rossiter

Charles E. Rossiter is Head of the Division of Computing and Statistics at the Clinical Research Centre, Harrow, UK.

Desert ecosystems

Arid-Land Ecosystems: Structure, Functioning and Management. Volume 1. Edited by D.W. Goodall, R.A. Perry and K.M.W. Howes. Pp. 881. (Cambridge University Press: Cambridge, 1979.) £45.

THIS volume is one of two that will give an overall account of arid-land studies, one of the themes of the International Biological Programme. Although much of the official IBP work was carried out in the USA, a considerable amount of similar work was also carried out in other countries at the same time, and much of this information has also been included. Unfortunately, and in spite of efforts by the editors, it has not proved possible to provide global coverage of arid regions. There is no account of South American deserts and the Sudan, a country with one of the largest desert areas, is only mentioned in passing. However, the areas that are described include North America, northern Africa, southern Africa, Australia, and South-West and Central Asia. The climate, soils, geomorphology, hydrology, flora and fauna are covered and the processes that operate within, and control, the ecosystem are dealt with under component, atmospheric, soil, plant and animal processes.

The editors' task must have been a nightmare for in spite of international cooperation investigators still seem to have operated within their own sub-set of published work and conventions, often set by linguistic or old colonial spheres of

influence. This can give rise to reliance on a very distorted range of literature or undue emphasis on a personal field of interest in a review article. Thus, one regional account emphasises productivity, another phytosociology, and in some accounts animals receive scant attention. The characterisation of regional climate is similarly diversely treated. In one account simplified 'Walter Klimadiagramm' are given, another resorts to rainfall tables, a third makes use of maps with isohyets, and others give the 'Walter Klimadiagramm' in full. Standardisation of representation would have allowed comparison and integration to be made more easily.

The part of the volume devoted to processes is more satisfactory and includes an excellent chapter on diversity and niche structure in desert communities by E.R. Pianka. Chapters on radiation, precipitation and soil processes deal well with specifically desert conditions. Other chapters treat topics more generally.

This book has not managed to distil the knowledge accumulated in the IBP studies on arid ecosystem structure and function in a unified way. It would be unfair, however, to be too critical as it is difficult to give constructive suggestions on how a more satisfactory account could have been achieved. It is perhaps best to accept the volume for what it is — a mammoth compendium of information on many of the world's desert ecosystems.

M. J. Chadwick

M. J. Chadwick is Senior Lecturer in Ecology at the University of York, UK.

Prosimian behaviour

The Study of Prosimian Behavior. Edited by G.A. Doyle and R.D. Martin. Pp. 696. (Academic: New York, San Francisco and London, 1979.) \$49.50.

The Study of Prosimian Behaviour is that rare pleasure: a book which fully achieves what it sets out to do. Its articles survey current knowledge of the prosimians, tabulate that knowledge by species and subject, and point out important hypotheses or conclusions for future study. From the opening chapter, in which Petter and Petter-Rousseaux give a revised classification, as well as the first full set of species distribution maps since Hill's tome *Primates: Comparative Anatomy and Taxonomy*, Vol.1 (Edinburgh University Press, 1953), it is clear that this is now the standard reference for most aspects of prosimian behaviour. As such, it belongs in any university library.

The editors are to be congratulated on assembling articles of fairly standard format, with a minimum of repetition,

although gestation lengths are given in three of the chapters and Martin and Charles-Dominique's basic decoding of home range and social structure in nocturnal forms enters at least three others. One might quibble that social structure of the diurnal species is under-represented, but this has been dealt with in recent books by A.F. Richard (*Behavioural Variation: Case Study of a Malagasy Lemur*; Bucknell University Press, Lewisburg; Associated University Presses, London, 1978 and edited by I. Tattersall and R.W. Sussman (*Lemur Biology*; Plenum, New York, 1975).

Chapters cover classification (J.J. Petter and A. Petter-Rousseaux), phylogenetic and allometric aspects of behaviour (R.D. Martin), reproduction (R.N. Van Horn and G.G. Eaton), maternal behaviour (P.H. Klopfer and K.J. Boskoff), development (G.A. Doyle), learning (B.J. Wilkerson and D.M. Rumbaugh), vocal communication (J.J. Petter and C.M. Hladik), diet (C.M. Hladik), spatial distribution (J.I. Pollock), vision (G. Pariente), olfaction (A. Schilling), locomotion (A. Walker), field studies of lorises (P. Charles-Dominique and S.K. Bearder),

and a study of *Tarsius bancanus* (C. Niemitz).

One theme which runs through the book is the importance of nocturnal life, where the prosimians may give us clues to our common ancestor. Charles-Dominique argues that small diurnal primates would compete with diurnal birds, which explains why small primates are nocturnal, and only larger ones diurnal. This size constraint then reappears in Doyle's and Martin's treatments of allometry in brain size and developmental rates.

The reader may trace his own interests. One such might be the "psychology of place". Wilson and Rumbaugh stress that prosimians approach formal learning tests by position, not by attention to objects. Then Charles-Dominique remarks that rainforest galagos and pottos can be trapped twice running if the trap is moved to a nearby site, for "It would seem the traumatic effect of capture is associated more with the place . . . then with the trap itself". Schilling then shows how scent-marking identifies places in both individual and social contexts — again, the recourse of a more nocturnal, less visually oriented evolutionary lineage. This, in turn, ties in with the views of Pollock, Bearder, and Charles-Dominique on ranging, "detour problems" for slow-climbing forms in the forest canopy, and social organisation by the overlapping of home ranges. For many prosimians, social life itself depends on position in the forest, cued by scent marks, rather than on vocal or visual contact with other animals.

One notable contribution is Georges Pariente's study of prosimian vision. Pariente presents photography of the living retina, electrophysiological studies, measurements of light intensity in Madagascar forest conditions, and activity of nocturnal species in relation to ambient light. He argues that "nocturnal life is more a shift in thresholds of sensitivity than a completely new, 'inversed' mode of existence. The notion of a night when vision is impossible only exists for strict diurnal species . . . Nocturnal prosimians can move around perfectly well in full daylight, using their vision in the usual way". He concludes that if the prosimian eye retains this degree of plasticity, one might propose a "primate ancestor which needed to see well by both day and night . . . small bodies, perhaps arboreal, with an appreciable amount of activity in the daytime . . . This ancestral form would have already been able to decode both in the intensity and frequency of light".

It is wholly appropriate that *The Study of Prosimian Behaviour* is dedicated to Georges Pariente, who was tragically killed in a traffic accident shortly after completing his article for this book.

Alison Jolly

Alison Jolly is Research Associate in the School of Biological Sciences at the University of Sussex, Brighton, UK.

announcements

Appointments

The Royal Society has awarded a Senior Research Fellowship to **Dr D.T. Edmonds**, university lecturer in physics in the University of Oxford and fellow of Wadham College, Oxford, for five years from 1 October 1979.

Amoz I. Chernoff, has been appointed Director of the Division of Blood Diseases and Resources in the National Heart, Lung, and Blood Institute.

Dr Aksel C Wiin-Nielsen (Denmark), has been appointed Secretary-General of the World Meteorological Organization (WMO) for a period of four years commencing 1 January 1980.

Geoffrey Caston, has been appointed Secretary General of the Committee of Vice-Chancellors and Principals in succession to **Sir Roy Marshall** who is taking up the Vice-Chancellorship of the University of Hull on 1 October 1979.

The following have been elected as Junior Beit Fellows as from 1 October 1979. **Farzin Farzaneh; Simon Brian Fishel; Virginia Jane Sandford Karlsson; Andrew John Parker; John Douglas Waterfield.**

Dr Jacob A. Brody, Associate Director, Epidemiology, Demography and Biometry Program, National Institute on Aging, National Institutes of Health, has been elected Vice-President of the American Epidemiological Society.

Dr Iain M. L. Donaldson has been appointed to the Chair of Applied Zoology in the University of Hull with effect from 1 October 1979.

Dr John W. Neale of the Department of Geology, University of Hull, appointed to a personal professorship with effect from 6 April 1979.

Dr Frank M Leslie, Reader in Mathematics in the University of Strathclyde since 1971, has been appointed to a Personal Professorship of Mathematics in the University.

Professor John Stuart, has been appointed Head of a new Department of Haematology in the University of Birmingham from 1 October 1979.

Robin Butlin has been appointed professor and head of Loughborough University's Department of Geography.

Neil W. Ashcroft, professor of physics at Cornell University, has been elected director of the University's Laboratory of Atomic and Solid State Physics for a five-year term with effect from 1 July 1979.

Dr Frank Wilkinson reader in physical chemistry at the University of East Anglia, has been appointed professor of physical chemistry in Loughborough University's Department of Chemistry.

Prof Michel Revel, Head of the Weizmann Institute's Virology Department is the first incumbent of the Ruth and Jerome A. Siegel and Freda and Edward M. Siegel Chair in Virology.

Dr R.F. Sellers, has been appointed Director of the Animal Virus Research Institute in succession to **Dr J.B. Brooksby**, who retires in December 1979.

Professor Bryan E. Richards has been appointed to the Mechan Chair of Aeronautics and Fluid Mechanics in the University of Glasgow, with effect from January 1980.

Dr Christopher Arme, Principal lecturer and Head of the Biology Unit at the North Staffordshire Polytechnic, has been appointed to a new Chair of Zoology in the Department of Biological Sciences at the University of Keele, from 1 October 1979.

Meetings

15-17 August, **20th Canadian High Polymer Forum**, Quebec (Dr Ian McEwan, Secretary Treasurer, 20th Canadian High Polymer Forum, Canadian Industries Limited, 1330 Castlefield Avenue, Toronto, Ontario, Canada).

21-24 August, **Cryogenic Engineering and International Cryogenic Materials Conferences**, Madison (Mrs Dee Belsher, CEC Administrator, National Bureau of Standards, Boulder, Colorado 80303).

9-14 September, **8th International Congress on Biometeorology**, Tel Aviv (World Meteorological Organization, Geneva, Switzerland).

13-15 September, **1979 Midwest Meeting of the American Geophysical Union**, Ohio Florida Avenue, N.W., Washington, D.C. 20009).

17-18 September, **26th Pacific Northwest Regional Meeting**, Oregon (American Geophysical Union, 2000 Florida Avenue, N.W., Washington, D.C. 20009).

19-26 September, **World Telecommunication Forum**, Geneva (Mrs Rita Glita, American Express Travel Service, 7 rue de Mont-Blanc, Boite postale 859, 1211 Geneve 1, Switzerland).

20 September, **The Oral Cavity: Disease and Drug Treatment**, London (Secretariat, c/o Institute of Biology, 41 Queen's Gate, London SW7, UK).

23-28 September, **A short course on Computer Ergonomics**, Loughborough (J.S. Wilcox, Administrative Assistant, Department of Human Sciences, University of Technology, Loughborough, Leicestershire, UK).

26-27 September, **584th meeting of the Biochemical Society**, Swansea (Meetings Officer, The Biochemical Society, 7 Warwick Court, High Holborn, London WC1, UK).

5 October, **The Nature of the Glassy State**, London (N. Rivier, Blackett Laboratory, Imperial College, London SW7, UK).

31 October, **5th Electronics Update Colloquium**, London (The Meetings Officer, The Institute of Physics, 47 Belgrave Square, London SW1, UK).

15 November, **Techniques of Teaching Crystallography**, London (The Meetings Officer, The Institute of Physics, 47 Belgrave Square, London SW1, UK).

3-7 December, **1979 Fall Meeting**, San Francisco (American Geophysical Union, 2000 Florida Avenue, N.W., Washington, D.C. 20009).

18 December, **Synthetic Developments in the Prostanoid Field**, London (Secretariat, c/o Institute of Biology, 41 Queen's Gate, London SW7, UK).

1980

2-4 January 1980, **17th Annual Solid State Physics Conference**, Coventry (The Meetings Officer, The Institute of Physics, 47 Belgrave Square, London SW1, UK).

24-26 January, **European Meeting on Intensive Care**, Paris (SOCFI, 7 rue Michel-Ange, F75016 Paris, France).

10-14 March, **1980 Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy**, Atlantic City (Frank W. Plankey, Department of Chemistry, University of Pittsburgh, Pennsylvania 15260).

17-18 March, **Modelling Dispersion of Transport Pollution**, Cambridge (The Institute of Mathematics and its Applications, Maitland House, Southend-on-Sea, Essex, UK).

18-20 March, **International Symposium on the Genetic Control of Natural Resistance to Infection and Malignancy**, Montreal (Dr Emil Skamene, Montreal General Hospital, Montreal, Quebec, Canada).

24-28 March, **Vacuum Devices Conference**, Cambridge (The Meetings Officer, The Institute of Physics, 47 Belgrave Square, London SW1, UK).

31 March-2 April, **3rd International Conference on Physicochemical Hydrodynamics**, Madrid (Prof. M.G. Velarde, Univ. Autonoma de Madrid, Cantoblanco (Madrid) Spain).

14-16 April, **Semi Insulating III-V Materials**, Nottingham (The Meetings Officer, The Institute of Physics, 47 Belgrave Square, London SW7, UK).

18-20 April, **1980 Annual Conference of the Education Group of the Institute of Physics**, Birmingham (The Meetings Officer, The Institute of Physics, 47 Belgrave Square, London SW1, UK).

26-27 April, **International Symposium on Bipolymers**, Birmingham (Dr H. Warson, 284 Warwick Road, Solihull, West Midlands, UK).

26-28 April, **6th Congress of German Cybernetics Society**, Mainz (Prof. I.W. v. Seelen, Institut für Zoologie, Universität Mainz, Saarstr. 21, D-6500 Mainz, FRG).

5-10 May, **1st International Conference on the Behaviour of Marine Animals, Ecology and Pollution**, Concarneau (Dr Yves Rouger, Marine Biology Laboratory, College de France, Concarneau 29100, France).

6-9 May, **9th World Congress on the Prevention of Occupational Accidents and Diseases**, Amsterdam (Benelux Organising Committee, Veiligheidsinstituut, Postbox 5665, 1007 AR Amsterdam, Netherlands).

27-30 May, **Joint meeting of the Institution of Mining and Metallurgy, the Society of Mining Engineers of AIME and the Metallurgical Society of AIME**, London (The Secretary, Institution of Mining and Metallurgy, 44 Portland Place, London W1, UK).

28-31 May, **IFAC Symposium on Water and Related Land Resource Systems**, Ohio (IFAC-Water Systems 1980, American Geophysical Union, 1909 K Street, N.W., Washington D.C. 20006).

9-14 June, **9th International Symposium on Organic Sulphur Chemistry**, Riga (The Organizing Committee, Institute of Organic Synthesis of the Latvian SSR Academy of Sciences, Aizkraukles 21, 226006 Riga, USSR).

10-12 June, **2nd International Conference on Developments in Power-System Protection**, London (The Press Office, Institution of Electrical Engineers, Savoy Place, London WC2, UK).

16-20 June, **9th International Coffee Science Symposium**, London (R.J. Clarke, General Foods Limited, Banbury, Oxon, UK).

24-26 June, **International Conference on Power System Monitoring and Control**, London (The Press Office, Institution of

Electrical Engineers, Savoy Place, London WC2, UK).

25 June, **Symposium on Trench and Fore-arc Sedimentation and Tectonics in Modern and Ancient Subduction Zones**, London (J.K. Leggett, Geological Society of London, Burlington House, London W1, UK).

29 June-6 July, **1980 International Palynological Congress on Diagenesis of Sedimentary Organic Matter**, Organic Preservation and Fossil Fuel Exploration, Cambridge (Dr Jim Brooks, Exploration Department, British National Oil Corporation, 150 St Vincent Street, Glasgow, UK).

21-26 June, **4th International Congress of Immunology**, Paris (Congres-Services, 1 rue Jules Lefebvre, 75009 Paris, France).

30 June-4 July, **13th International Symposium on Chromatography**, Cannes (G.A.M.S. 88 boulevard Malesherbes, 75008 Paris, France).

7-11 July, **2nd International Congress on Toxicology**, Brussels (SdR Associated, 16 Avenue de Abeilles, 1050 Brussels, Belgium).

7-12 July, **8th International Congress of Primatology**, Florence (B. Chiarelli, Institute of Anthropology, Via del Proconsolo 12, 50127 Firenze, Italy).

7-17 July, **26th International Geological Congress**, Paris (Secretariat General due 26eme Congrès Géologique International, Maison de la Géologie, 77-79 rue Claude Bernard, 75005 Paris, France.).

14-19 July, **9th International Conference on General Relativity and Gravitation**, Jena (Professor E. Schmutzer, GR9 Sektion Physik, Friedrich-Schiller-Universität, DDR-69 Jena, Max-Wien-Platz 1, FRG).

14-24 July, **3rd International Symposium on Water Rock Interaction**, Edmonton (Dr Brian Hitchon, Secretary-General, WRI-3, Alberta Research Council, 11315-87 Avenue, Edmonton, Alberta, Canada).

22-26 July, **4th International Conference on Cyclic Nucleotides**, Brussels (J.E. Dumont, I.R.I.B.H.N., School of Medicine, 2 rue Evers, B-1000 Brussels, Belgium).

4-9 August, **7th International Conference on Raman Spectroscopy**, Ottawa (Ken Charbonneau, Conference Services, National Research Council of Canada, Ottawa, Ontario, Canada).

2-5 September, **7th International Symposium on Medicinal Chemistry**, Madrid (7th International Symposium on Medicinal Chemistry, 142-144 Oxford Road, Cowley, Oxford, UK).

3-5 September, **Microbial Adhesion to Surfaces**, Reading (Dr P.R. Rutter, c/o The Conference Secretariat, Society of Chemical Industry, 14 Belgrave Square, London SW1, UK).

5-7 September, **17th Symposium of European Association against Virus Diseases**, Munich (Professor F. Deinhardt, Max-von-Pettenkofer-Institut, Pettenkoferstrabe 9a, D-8000 München 2, FRG).

Reports & publications UK & Ireland — June

Air Pollution — Its Dispersion and Effects. Pp. 59. £4. Water Pollution — Its Dispersion and Effects. Pp. 49. £4. (Atkins Research and Development Monographs.) (Epsom, Surrey: Atkins Research and Development, Woodcote Grove, Ashley Road, 1979.) [46]

Health Service Commissioner. First Report for Session 1979-80: Investigations completed — December 1978 to March 1979. Pp. 216. (London: HMSO, 1979.) £3.25 net. [66]

Health and Safety Executive. Isocyanates: Toxic Hazards and Precautions. (Guidance Note EH 16). Pp. 8. (London: HMSO, 1979.) 38p plus postage. [66]

Review Body on Doctors' and Dentists' Remuneration — Ninth Report, 1979. Chairman: Sir Ernest Woodroffe. Pp. v + 127. (Cmnd. 7574). (London: HMSO, 1979.) £2.25 net. [76]

Department of the Environment. Nineteenth Progress Report of the Standing Technical Committee on Synthetic Detergents. Pp. iv + 16. (London: HMSO, 1979.) 75p net. [86]

Clinical and Laboratory Haematology, Vol. 1, No. 1, 1979. Edited by J.M. England. Pp. 1-73. Published quarterly. Annual subscription: £21 (UK), £26 (overseas), and \$57-50 (USA and Canada). (Oxford and London: Blackwell Scientific Publications, 1979.) [86]

Natural Environment Research Council: Institute of Terrestrial Ecology. Sampling. By J.N.R. Jeffers. (Statistical Checklist, 2). Pp. 7. (Cambridge: Institute of Terrestrial Ecology, 68 Hills Road, 1979.) [116]

Freshwater Biological Association. 1929-1979: The First Fifty Years. By G.E. Fogg. Pp. 39 (Ambleside: Freshwater Biological Association, 1979.) [116]

The Wellcome Trust, 1976-78: Twelfth Report. Pp. 113. (London: The Wellcome Trust, 1 Park Square West, 1979.) [116]

Council for Science and Society, The Outer Circle Policy Unit. The Big Public Inquiry. (A Proposed New Procedure for the Impartial Investigation of Projects with Major National Implications.) Chairman of Working Party: Paul Sieghart. Pp. iii + 72. (London: The Outer Circle Policy Unit, 4 Cambridge Terrace, NW1, 1979.) £1.50. [116]

Health and Safety Executive. Publication Catalogue '79. Pp. 107. (London: HMSO, 1979.) £1 net. [136]

Handbook of British Medical Research Charities. Pp. 35. (London: The Association of Medical Research Charities, c/o British Heart Foundation, 57 Gloucester Place, W1, 1979.) [146]

The Rowett Research Institute. Annual Report of Studies in Animal Nutrition and Allied Sciences, Vol. 34, 1978. Pp. 131. (Bucksburn, Aberdeen: The Rowett Research Institute, 1979.) £2.50. [156]

Propolis: a Review by E.L. Ghisalberti. (*Bee World* 60(2), 1979). Pp. 26. (Gerrards Cross, Bucks.: International Bee Research Association, Hill House, 1979.) 60p or \$1.40, plus 10p or \$0.25 p&p. [196]

The Radiochemical Centre. Technical Bulletin 79/4: Radiolabelled Neurochemicals for Receptor Studies. Pp. 8. (Amersham: The Radiochemical Centre, 1979.) [196]

International Conference on Recombination in Semiconductors. (*Solid-State Electronics: An International Journal*, Vol. 21, Nos. 11/12, November/December 1978.) Pp. vi + 1273-1612. (Oxford and New York: Pergamon Press, 1978.) £20. [196]

Proceedings of The 1978 Canadian Reliability Symposium, October 19-20, 1978, Ottawa, Ontario, Canada. Chairman: Nicholas Balke. (*Microelectronics and Reliability: An International Journal and World Abstracting Service*, Vol. 19, Nos. 1/2, 1979.) Pp. 1-168. (Oxford and New York: Pergamon Press, 1979.) £20. [196]

Catalogue of the Ocean Bottom Deposits Collection in the British Museum (Natural History). Part 1: Atlantic and Arctic Oceans including the European Seas. By H. A. Buckley, C. J. Elliott, Nancy M. Graham, L. R. Johnson, D. R. C. Kempe and D. B. Williams. On Microfiche. Pp. 12. (London: British Museum (Natural History), 1979.) £12. [206]

Countryside Commission. Recreational Paths. Pp. 24. (Cheltenham: Countryside Commission, John Dower House, Crescent Place, 1979.) gratis. [206]

Cancer Research Campaign, (British Empire Cancer Campaign for Research). 56th Annual Report, 1978. Pp. 171. (London: Cancer Research Campaign, 2 Carlton House Terrace, SW1, 1979.) [206]

Grassland Research Institute. Annual Report 1978. Pp. xii + 133. (Hurley, Maidenhead: The Grassland Research Institute, 1979.) £2.50 [216]

Health and Safety Executive. Safety During Semiconducting and Continuous Casting of Copper and Copper-Based Alloys. Pp. 18. 80p net. Molten Metal and Water Explosions. Pp. 24. £1 net. (Joint Standing Committee on Health, Safety and Welfare in Foundries.) (London: HMSO, 1979.) [216]

— July

Royal Observatory, Edinburgh. Annual Report, 1977-1978. Pp. 54. (Edinburgh: Royal Observatory, 1979.) [97]

Annual Report of the Health Service Commissioner 1978-79. Pp. 33. (London: HMSO, 1979.) 80p net. [97]

United Kingdom Atomic Energy Authority. Radioactive Fallout in Air and Rain: Results to end of 1978. By R.S. Cambray, Miss E.M.R. Fisher, K. Playford, J.D. Eakins and D.H. Peirson. Pp. 43. (London: HMSO, 1979.) £1.50 net. [97]

The Macaulay Institute for Soil Research. 1977-1978 Annual Report. No. 48. Pp. 139. (Craigiebuckler, Aberdeen: The Macaulay Institute for Soil Research, 1979.) [97]

John Innes Institute. Sixty Ninth Annual Report, 1978. Pp. 151. (Norwich: John Innes Institute, 1979.) £3. [97]

List of University Institutions in the Commonwealth. 23rd edition. Pp. 37. (London: The Association of Commonwealth Universities, 1979.) [97]

Ancient Monuments Board for England. Twenty-fifth Annual Report, 1978. Pp. 40. (London: HMSO, 1979.) £1.50 net. [97]

The British Industrial Biological Research Association. Annual Report, 1978. Pp. 35. (Carshalton, Surrey: BIBRA, BIBRA, 1979.) [97]